Molecular Basis of *Escherichia coli* L/F transferase: Catalytic Mechanism and Substrate Specificities

by

Angela Wai Shan Fung

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Biochemistry
University of Alberta

Abstract

The *Escherichia coli* leucyl/phenylalanyl-tRNA protein transferase (L/F transferase) catalyzes the tRNA-dependent post-translational addition of amino acids onto the N-terminus of a protein polypeptide substrate. The enzymatic N-terminal addition of an amino acid to a protein has been identified as a molecular marker to target proteins for degradation via the N-end rule pathway, where it determines the relationship between the *in vivo* half-life of a protein and the identity of its N-terminal amino acid.

Here we investigate the molecular basis of the catalytic mechanism, substrate analogue design and tRNA substrate recognition by L/F transferase through the analysis of available X-ray crystal structures, mutagenesis, *in vitro* transcribed tRNAs, and an enzyme functional assay that was developed by our lab (quantitative matrix-assisted laser desorption/ionization time of flight mass spectrometry assay).

The N-terminal post-translational addition of an amino acid is analogous to that of peptide bond formation by the ribosome. A previous protein-based catalytic mechanism for L/F transferase has been proposed. Our study on the functional role of D186, a proposed catalytic residue, illustrates that D186's function is to orient substrates. We propose an alternative substrate-assisted proton shuttling catalytic mechanism, similar to one proposed for the ribosome.

The molecular details on tRNA recognition by L/F transferase is studied using crystal structures with tRNA substrate analogues bound,

despite differences in their binding orientations. We investigate and illustrate that this difference is due to the different modifications on the analogues. This study leads to the first steps to the design and development of improved substrate analogues for this class of enzyme.

Contrary to previous studies suggesting that L/F transferase mainly recognizes the 3' aminoacyl adenosine of an aminoacyl-tRNA for substrate recognition, our studies shed light on the critical importance of recognition of the remaining tRNA body especially the acceptor stem in a sequence-dependent manner.

Taken together, our molecular studies into the L/F transferase reaction expand the current understanding of the molecular details of the catalytic mechanism, substrate analogue design, and tRNA substrate recognition for L/F transferase.

Preface

This thesis is an original work by Angela Wai Shan Fung (AWSF). As the primary author for the following manuscripts, AWSF prepared a majority of the materials, executed and analyzed all experiments. AWSF and RPF designed the experiments and prepared the following manuscripts.

Chapter 1 and Chapter 6 of this thesis have been submitted for publication as AW Fung and RP Fahlman (2014) The Molecular Basis of Post-Translational Addition of Amino Acids in *E. coli* L/F transferase.

Chapter 3 of this thesis has been published as AW Fung, HA Ebhardt, H Abeysundara, J Moore, Z Xu, and RP Fahlman (2011) An Alternative Mechanism for the Catalysis of Peptide Bond Formation by L/F Transferase: Substrate Binding and Orientation. *Journal of Molecular Biology.* 409 (4): 617-629. HAE contributed to discussions on the early design of the experiment. HA, JM, and ZX contributed to the preparation of materials.

Chapter 4 of this thesis has been published as AW Fung, HA Ebhardt, KS Krishnakumar, J Moore, Z Xu, P Strazewski, and RP Fahlman (2014) Probing the Leucyl/Phenylalanyl tRNA Protein Transferase Active Site with tRNA Substrate Analogues. *Protein and Peptide Letters*. 21 (7): 603-614. HAE contributed to discussions on the early design of the experiment. KSK, JM, ZX and PS contributed to the preparation of materials.

Chapter 5 of this thesis has been published as AW Fung, CC Leung and RP Fahlman (2014) The Determination of $tRNA^{Leu}$ Recognition Nucleotides for *Escherichia coli* L/F transferase. *RNA*. 20 (8): 1210-1222. CCL contributed to the preparation of the docking model figure presented.

“Sometimes the questions are complicated and the answers are simple.”

-- Dr. Seuss

Acknowledgements

First and foremost I would like to thank my supervisor - Dr. Richard Fahlman - for his trust, support and guidance. Thank you for being patient with me and giving me the freedom to explore my research interests. Richard is a brilliant and diligent scientist who has been a great mentor and someone I will undoubtedly try to model myself after in the future. Richard not only enlightens me with scientific insights, but also inspires me the way of life. Thank you for being my mentor, friend, and family.

My research on tRNA substrate analogue binding has been a collaboration with Dr. Kollappillil Krishnakumar and Dr. Peter Strazewski at the Université de Lyon. My committee members, Dr. Mark Glover and Dr. Luc Berthiaume, have been very supportive and encouraging during my graduate studies. A special thanks to my internal examiner, Dr. George Owttrim, and my external examiner, Dr. Ute Kothe, for their support and insights. I would like to thank NSERC, AITF, and the University of Alberta for their scholarship support.

I would also like to thank past and current members of the Fahlman lab, especially Dana, Susan, Alex, Steve, Roshani, Prarthna, Julian, Nancy, Ben, Les, Mohamed, Guru, and more for providing their support. Special thanks to the summer/499 students that I have supervised, Heshani, David, Aaron, Devin, and Michal, who have contributed to parts of this research. Thanks to Jack and Paul from IBD for teaching me mass spectrometry fundamentals. It is my greatest pleasure to have worked (and partied) with these amazing and fun people.

Graduate school has encompassed some of the greatest times of my life, a period that I will always cherish. Special thanks to Stephen for being an inspirational and supportive friend, I will always remember you. Thanks to all the friends from the biochemistry department, particularly Ply, Andrew, Justin, Curtis, Ross, Tamara, Jen, Pam, Grant, Daniel and more. They have made staying and hanging out in the department fun. Thanks to all the friends and colleagues from STRG and RiboWest.

Special thanks to my friends who have supported and being there for me throughout my graduate study, especially Daniel, Edmond, Gloria, Haley, Natasha, Zohadya, Nisa, Mihiri, Michael, Clement, Linda, Claudia, Kenneth, Elysia, Kaiser, and Hazel.

A very special thanks and love to Charles. He has supported and encouraged every step of the way, even when it meant sacrifice on his part. Charles, I love you and thank you for supporting me to grow and be better every day. Thanks to Charles' family, who have been kind and supportive. Finally and most importantly, thanks to Mom, Dad, Kingsley and my family for their unconditional love and continuous support. I love you all. Hugs!

Table of Contents Page

List of Tables Page

List of Figures Page

List of Abbreviations

Φ	hydrophobic amino acids
π	small aliphatic amino acids
a_1	acceptor peptide binding pocket for 1st position
a_2	acceptor peptide binding pocket for 2nd position
aa-RS	aminoacyl-tRNA synthetase
aat	aminoacyl transferase (gene encoding L/F transferase)
aa-transferase	aminoacyl-tRNA protein transferase
aa-tRNA	aminoacyl-tRNA
AAA+	ATPases associated with various cellular activities
Ac/N-end rule	eukaryotic N-terminal acetylated N-end rule
Arg/N-end rule	eukaryotic arginine N-end rule
ATE1	arginine transfer enzyme 1
ATEL1	ATE-like 1
ATP	adenosine triphosphate
Bpt	bacterial protein transferase
BSA	bovine serum albumin
ClpA	unfoldase/ translocase, chaperone for ClpP
ClpP	serine protease
ClpS	adaptor protein for ClpA

cydCD	gene encoding cysteine ABC transporter ATPase
d_1	donor substrate binding pocket 1
Dps	DNA protection during starvation protein
Dupli-GNAT	duplicated GNAT protein superfamily
E. coli	*Escherichia coli*
EF-Ts	elongation factor thermal stable
EF-Tu	elongation factor thermal unstable
Fem	factors essential for methicillin
fMet	*N*-formylmethionine
GDP	guanosine diphosphate
GFP	green fluorescent protein
GNAT	GCN5-related N-acetyltransferase
GTP	guanosine triphosphate
K_a	association constant
k_{cat}	turnover number
k_{cat} / K_M	catalytic efficiency
K_i	inhibitiory constant
K_M	Michaelis-Menten constant
Leu-tRNA^{Leu}	leucyl-tRNA (aminoacylated or charged)
L/F transferase	leucyl/phenylalanyl tRNA protein transferase
LC-MS/MS	liquid chromatography tandem mass spectrometry

MALDI-ToF	matrix-assisted laser desorption ionization time of flight
MetAP	methionine aminopeptidase
mRNA	messenger RNA
MS	mass spectrometry
mprF	gene for multipeptide resistance factor
N-degron	N-terminal degradation signal
N-recognin	N-terminal recognition signal
NBD	nitrobenzoxadiazole
Ni^{2+}-NTA	nickel (II) nitrilotriacetic acid
PATase	putrescine aminotransferase
PDF	peptide deformylase
PG	phosphatidylglycerol
pK_a	acid dissociation constant at logarithmic scale
PPase	inorganic pyrophosphatase
PP_i	inorganic pyrophosphate
P_i	inorganic phosphate
PutA	proline utilization protein A
rA-aa	aminoacyl ribonucleoadenosine
rA-Phe	phenylalanyl ribonucleoadenosine
RNA	ribonucleic acid

TLC	thin layer chromatography
tRNA	transfer ribonucleic acid
$tRNA^{Leu}$	leucine-tRNA (deacylated or uncharged)
UV	ultra violet

Note:

[1] Amino acid residues of a protein will be written in the one-letter code followed by the residue number (ex. D186)

[2] N-terminal amino acid specificity or aminoacyl moiety from an aminoacyl-tRNA will be written in the three-letter code (ex. Met)

[3] tRNA nucleotides numbering will be written as subscripts (ex. A_{76})

Chapter 1

Introduction

A version of this chapter is submitted:

Fung AW and Fahlman RP (2014) The Molecular Basis of Post-Translational Addition of Amino Acids in *Escherichia coli* N-end Rule Pathway. Submitted.

1.1. Intracellular Proteolysis and *E. coli* N-end Rule

Bacteria live in a dynamic environment where temperature, pH, nutrients, and chemical agents may vary drastically. These changes in the environment can result in cellular stress and potentially result in protein misfolding and the accumulation of insoluble aggregates. Intracellular proteolysis has been regarded as the method to clear damaged or misfolded proteins from a cell. More recently it has also been implicated to regulate cellular processes in bacteria such as protein quality control (Dougan *et al.* 2002a), pathogenesis (Butler *et al.* 2006), and development (Jenal 2009, Moliere and Turgay 2009). Regulated intracellular proteolysis in *E. coli* is carried out by processive ATP-dependent proteases including ClpAP, ClpXP, Lon, FtsH, and HslUV (Gur *et al.* 2011). Each protease degrades a subset of substrates depending on specific expression patterns, subcellular localization, and substrate specificity (Gur *et al.* 2011, Sauer and Baker 2011). For example, Lon selectively recognizes misfolded proteins whereas FtsH primarily recognizes membrane-bound substrates (Dougan *et al.* 2002a, Gur *et al.* 2011, Sauer and Baker 2011).

As intracellular proteolysis is an energy-dependent and irreversible process, tight regulation on substrate recognition is essential to avoid random, unscheduled degradation. The degradation signal (or degron) that targets the protein for degradation varies. Some degrons expose upon an external stimulus or stress, while some degrons are short

sequences located in the unstructured regions within the N- or C-terminus. One simple degradation signal is the “N-degron” based on the N-end rule, where the identity of the N-terminal amino acid dictates the *in vivo* half-life of a protein (Bachmair *et al.* 1986). In *E. coli*, N-terminal amino acid residues are categorized into stabilizing, primary (1º) destabilizing and secondary (2º) destabilizing residues.

Variations of the N-end rule pathway are widely present in eubacteria (Tobias *et al.* 1991, Ichetovkin *et al.* 1997), fungi (Bachmair and Varshavsky 1989), plants (Potuschak *et al.* 1998, Graciet and Wellmer 2010) and mammals (Gonda *et al.* 1989). In eukaryotes, there are currently two major branches of the N-end rule pathway (See review (Gibbs *et al.* 2014)). The eukaryotic Arg/N-end rule pathway involves unacetylated, free amino N-terminal destabilizing residues (ex. Leu, Arg etc.) that are generated or modified by enzymes including caspases, calpains, other non-processive proteases, N-terminal amidases, and R-transferases (Bachmair *et al.* 1986). The eukaryotic Ac/N-end rule pathway involves acetylated N-terminal destabilizing residues (ex. Ac-Ala, Ac-Ser etc.) generated or modified by enzymes such as methionine aminopeptidases and N-terminal acetylases (Hwang *et al.* 2010). Recently it has also been identified that the previously stable N-terminal Met is also a 1º destabilizing residue under the condition that it is followed by a hydrophobic residue (Φ) such that N-terminal MetΦ can participate in both Arg/N-end rule and Ac/N-end rule pathway depending on its

acetylation state (Kim *et al.* 2014). Thus the N-degron may be generated by co-translational, post-translational, or conditional non-processive proteolytic cleavages and modifications. Subsequently eukaryotic N-degrons are recognized and degraded by the ubiquitin-proteasome systems.

The *E. coli* Leu/N-end rule pathway is similar to the eukaryotic Arg/N-end rule pathway organization and recognizes unacetylated, free amino N-termini. *E. coli* proteins with N-terminal 1º and 2º destabilizing residues have an *in vivo* half-life of approximately 2 minutes, while the remaining amino acid residues are considered stable with an *in vivo* half-life of over 10 hours (Tobias *et al.* 1991). In *E. coli*, 1º destabilizing residues include Leu, Phe, Tyr, and Trp, while 2º destabilizing residues include Arg and Lys. **Figure 1-1** illustrates the major steps in the *E. coli* Leu/N-end rule pathway: 1) potential proteolytic cleavage by unknown non-processive protease, 2) post-translation addition of amino acid modification by L/F transferase, and 3) recognition by ClpS and degradation by ClpAP.

1.1.1. Non-Processive Proteolytic Cleavage

All proteins synthesized in bacteria begin with a stable N-terminal formyl-methionine (fMet). **Figure 1-2** illustrates some possible mechanisms of non-processive proteolytic cleavage that expose N-terminal destabilizing residues for N-end rule-mediated degradation. Peptide deformylase (PDF) and methionine aminopeptidase (MetAP) are

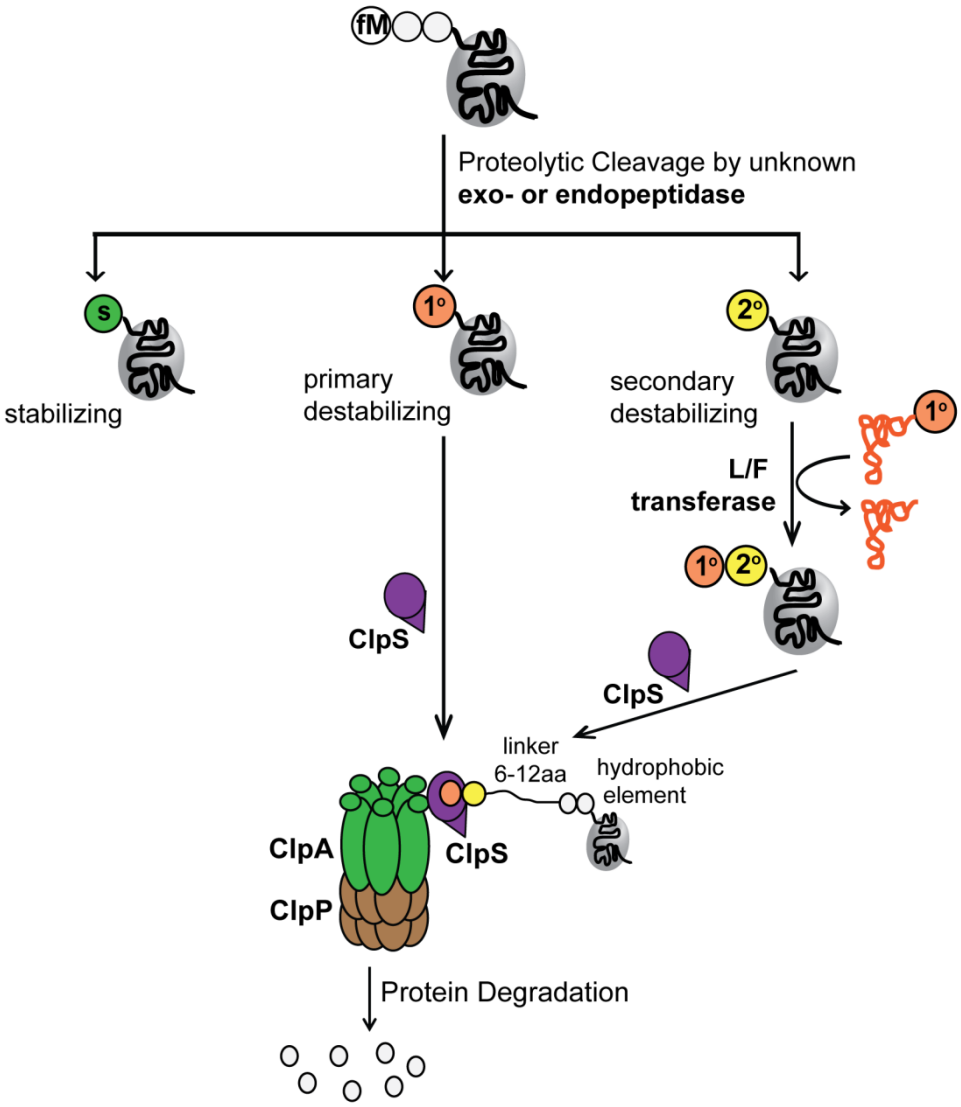

Figure 1-1: A schematic of *E. coli* N-end rule pathway. All *E. coli* proteins initiate with formyl-methionine (fM). Upon proteolytic cleavage by yet unknown exo- or endopeptidase a neo-N-terminus is exposed, which is categorized into stabilizing (s, green), primary destabilizing (1º, pink, including Leu, Phe, Trp, and Tyr) or secondary destabilizing (2º, yellow, including Arg and Lys). Secondary destabilizing N-degrons are recognized by L/F transferase, which post-translationally adds a Leu or Phe (1º destabilizing, pink) from its cognate aminoacyl-tRNA to the N-terminus of a protein bearing a Arg or Lys (2º destabilizing, yellow). The adaptor protein ClpS (purple) recognizes and binds to proteins with a 1º destabilizing N-degron and delivers the tagged protein to the proteasome-like complex ClpAP, where the protein is unfolded by the AAA+ ATPase ClpA (green) and degraded by the protease ClpP (brown).

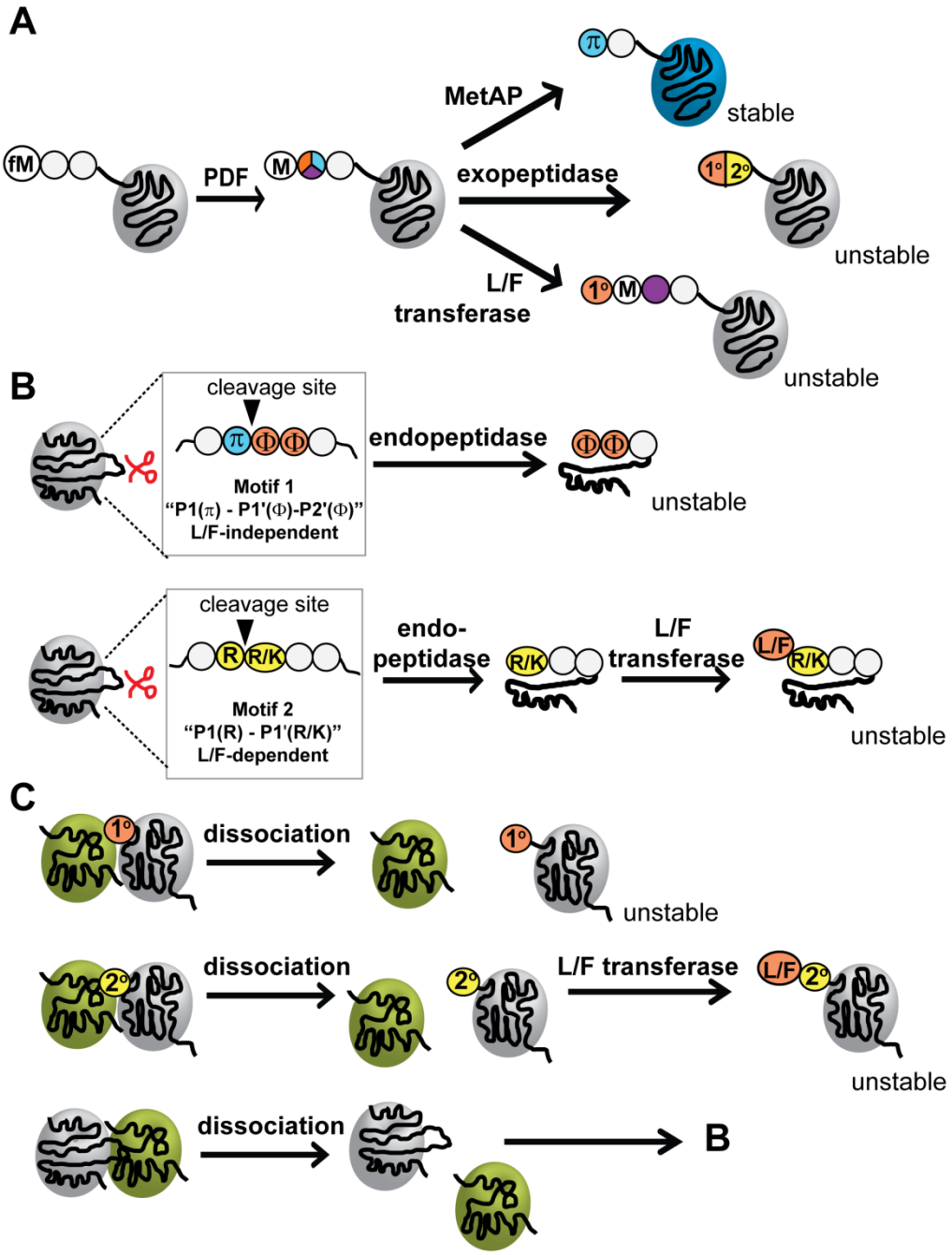
A
fM
PDF
M
MetAP
π
stable
exopeptidase
1° 2°
unstable
L/F
transferase
1° M
unstable
B
cleavage site
π Φ Φ
Motif 1
"P1(π) - P1'(Φ)-P2'(Φ)"
L/F-independent
endopeptidase
Φ Φ
unstable
cleavage site
R R/K
Motif 2
"P1(R) - P1'(R/K)"
L/F-dependent
endo-
peptidase
R/K
L/F
transferase
L/F R/K
unstable
C
1°
dissociation
1°
unstable
2°
dissociation
2°
L/F transferase
L/F 2°
unstable
dissociation
B

Figure 1-2: A proposed mechanism for the generation of N-end rule substrates. A) All *E. coli* proteins synthesized by the ribosome initiates with formyl-methionine (fM). N-terminal processing by peptidyl deformylase (PDF) removes the formyl group from methionine (M). N-terminal methionine can be further modified by (**top**) methionine aminopeptidases (MetAP) when the second amino acids are small (π), (**middle**) unknown exopeptidase that exposes N-degron (1º = pink and 2º = yellow), or (**bottom**) L/F transferase when the second amino acid offers favourable binding at the a_2 site. **B)** Proposed motifs of (**top**) L/F transferase-independent and (**bottom**) L/F transferase-dependent proteolytic cleavage sites on the surface, unstructured loops by yet to be identified endopeptidases. The L/F transferase-independent motif 1 "P1(π) – P1'(Φ)-P2'(Φ)" has a cleavage site between the small aliphatic amino acid (π, ex. Gly and Ala) at position P1 and the hydrophobic amino acid (Φ) at position P1', exposing 1º N-degron (pink including Leu, Phe, Trp, and Tyr). The L/F transferase-dependent motif 2 "P1(R) – P1'(R/K)" has a cleavage site between Arg at position P1 and a basic residue at position P1', exposing the 2º N-degron (yellow including R and K) for further modification by L/F transferase. **C)** The dissociation of large protein complex subunits exposes (**top**) 1º N-degron (pink), (**middle**) 2º N-degron for L/F transferase-dependent modification (yellow), or (**bottom**) pro-N-degron for further proteolytic cleavages. The use of symbols for amino acids is according to (Aasland *et al.* 2002).

two enzymes known to modify the N-terminus (**Figure 1-2A**) (Dougan *et al.* 2010). PDF hydrolyzes the removal of the *N*-formyl group from N-terminal fMet to form N-terminal Met, which subsequently allows MetAP to hydrolyze the removal of N-terminal Met when adjacent to a small aliphatic residue (π, ex. Gly, Ala) (Dougan *et al.* 2010). The resulting new N-terminus by PDF and MetAP processing are stabilizing residues and thus are stable. Since the 1º and 2º destabilizing residues are large hydrophobic or charged residues, PDF and MetAP are not likely to be the enzymes responsible for the generation of N-degron. Alternatively, proteolytic cleavage by yet unknown exopeptidase from the N-terminus may expose 1º or 2º destabilizing N-degrons under specific conditions (Dougan *et al.* 2010, Humbard *et al.* 2013). Recently, the identification of the first L/F transferase substrate, putrescine aminotransferase (PATase), demonstrates that L/F transferase post-translationally add a Leu or Phe (1º destabilizing) to the N-terminus of proteins with an N-terminal Met residue (stable) targeting the modified protein for degradation (Ninnis *et al.* 2009). This may serve as another mechanism for the generation of N-degrons.

In addition to N-terminal processing and modifications, another method for exposing a new N-terminus is the use of non-processive endopeptidases cleaving exposed, unstructured regions of a protein (**Figure 1-2B**) (Dougan *et al.* 2010, Humbard *et al.* 2013). This proposed non-processive proteolytic cleavage by an endopeptidase to generate an

N-end rule substrate has been observed in yeast (Rao *et al.* 2001). L/F transferase-independent and –dependent endopeptidase motifs have been proposed based on the proteome-wide identification of *in vivo E. coli* N-end rule substrates (Humbard *et al.* 2013). The L/F transferase-independent pro-N-degron motif is "P1(π) – P1'(Φ)-P2'(Φ)" where the cleavage occurs between the small, aliphatic amino acid (π, ex. Gly, Ala) at position P1 and hydrophobic amino acid (Φ, ex. Leu, Phe) at position P1'. The exposed new N-terminal hydrophobic residue at P1' position is a 1º destabilizing and is recognized and degraded. The L/F transferase-dependent pro-N-degron motif is "P1(R) – P1'(R/K)", where the cleavage occurs between Arg at position P1 and a basic residue at position P1'. Upon the exposure of basic 2º destabilizing residue at P1' position, L/F transferase can further modify and target proteins for degradation. Thus, there are at least two classes of endopeptidase with different cleavage specificity based on the pro-N-degron motifs observed. Alternatively, the dissociation of subunits from large complexes may expose 1º or 2º N-degron or cleavage sites for the specific targeting of these proteins for degradation (**Figure 1-2C**).

The specific mechanism that generates *E. coli* N-end rule substrates will remain elusive until the identification of proteases that are responsible. It is commonly observed that the degradation signals on proteins become exposed upon stress induction (Gur *et al.* 2011). We hypothesize that *E. coli* N-end rule and peptidases/proteases involved are

initiated and activated under a stress or metabolic response. It will be important to determine whether the N-end rule in *E. coli* regulates enzymes in distinct pathways when triggered by general or specific stresses such as nutrition, temperature, salt, metal, pH changes and more.

1.1.2. Post-Translational Addition of Amino Acid Modification

Alternative to non-processive proteolytic cleavage, a protein can become destabilized by the addition of destabilizing amino acids to the N-terminus of proteins. Aminoacyl-tRNA protein transferases (aa-transferases) are a class of enzymes that catalyzes tRNA-dependent post-translational addition of amino acids from an aa-tRNA to the N-terminus of a protein (Soffer *et al.* 1969). The post-translational addition of amino acids to proteins in eukaryotes has been demonstrated to have a variety of physiological functions *in vivo,* such as heart development (Kwon *et al.* 2002), G-protein signalling (Hu *et al.* 2005, Lee *et al.* 2005), neurodegeneration (Bongiovanni *et al.* 1995), gametogenesis (Leu *et al.* 2009), and apoptosis (Ditzel *et al.* 2003, Herman-Bachinsky *et al.* 2007, Wickliffe *et al.* 2008) (see review for more details (Saha and Kashina 2011)). Despite an emerging understanding of the plethora of physiological functions for the post-translational addition of amino acids in eukaryotes, there is only a minimal understanding on the molecular basis of structure, catalytic mechanism, substrate specificity, and regulation of the eukaryotic aa-transferase ATE1. The ~60 kDa ATE1 ($R^{C*,D,E}$

transferase; E.C. 2.3.2.8, gene *ate1* for arginine transfer enzyme 1), catalyzes the transfer of Arg (1º destabilizing) from Arg-tRNA^{Arg} to the N-terminus of proteins having an Asp, Glu, or oxidized Cys (2º destabilizing) (Gonda *et al.* 1989). ATE1 is involved in the Arg/N-end rule pathway and decreases the stability of the modified proteins (Bachmair *et al.* 1986). In mammals the biological importance of ATE1 is highlighted by the observation that *ate1*$^{-/-}$ mice are embryonic lethal, exhibiting impaired cardiac development and other malformations (Kwon *et al.* 2002).

Meanwhile the physiological function of N-end rule in prokaryotes has been more elusive since only a couple of substrates have been identified suggesting roles in putrescine homeostasis, proline catabolism and growth phase-dependent proteolysis (Ninnis *et al.* 2009, Schmidt *et al.* 2009). However the X-ray structural data and biochemical techniques available for investigating prokaryotic aa-transferase make them ideal surrogates for understanding the molecular mechanisms of this class of enzymes (Suto *et al.* 2006, Watanabe *et al.* 2007, Ebhardt *et al.* 2009, Wagner *et al.* 2011).

In *E. coli*, the sole aa-transferase is leucyl/phenylalanyl tRNA protein transferase (L/F^{R,K} transferase, E.C. 2.3.2.6, gene *aat* for aminoacyl transferase) (Leibowitz and Soffer 1969, Shrader *et al.* 1993). *In vitro*, L/F transferase catalyzes the transfer of a Leu or Phe (1º destabilizing) or Met (stabilizing, although to a lesser extent) from an aa-tRNA to the N-terminus of a substrate protein bearing an N-terminal Arg or

Lys (2º destabilizing) or Met (stabilizing) residue (Soffer 1974, Scarpulla *et al.* 1976, Abramochkin and Shrader 1996, Ninnis *et al.* 2009) (**Figure 1-3**). It has been shown that the dominant modification is leucylation *in vivo*, hence the naming of Leu/N-end rule (Shrader *et al.* 1993).

This thesis will focus on the molecular investigations on the catalytic mechanism, substrate specificities, and regulation of L/F transferase using available X-ray crystal structures and a more sensitive functional assay developed by our laboratory (Ebhardt *et al.* 2009, Fung *et al.* 2011, Fung *et al.* 2014a, Fung *et al.* 2014b).

1.1.3. Recognition and Degradation by ClpSAP

Considerable advances in the area of *E. coli* N-end rule recognition and degradation have been investigated (see reviews for more details (Dougan *et al.* 2010, Dougan *et al.* 2012)). In *E. coli*, the 1º destabilizing residues are recognized and delivered to the ClpAP protease complex by the adaptor protein ClpS (N-terminal recognition, N-recognin) (Guo *et al.* 2002, Zeth *et al.* 2002, Erbse *et al.* 2006, Wang *et al.* 2008b, Schuenemann *et al.* 2009). With its bound substrate, ClpS docks to and triggers a conformational change in the unfoldase/ translocase ClpA and the substrate is transferred from ClpS to ClpA (Guo *et al.* 2002, Zeth *et al.* 2002, Erbse *et al.* 2006, Wang *et al.* 2008b, Schuenemann *et al.* 2009). ClpA, a member of the Clp/Hsp100 family of molecular chaperones, contains two AAA+ (ATPases associated with various cellular activities) domains and is capable of unfolding and translocating protein substrates

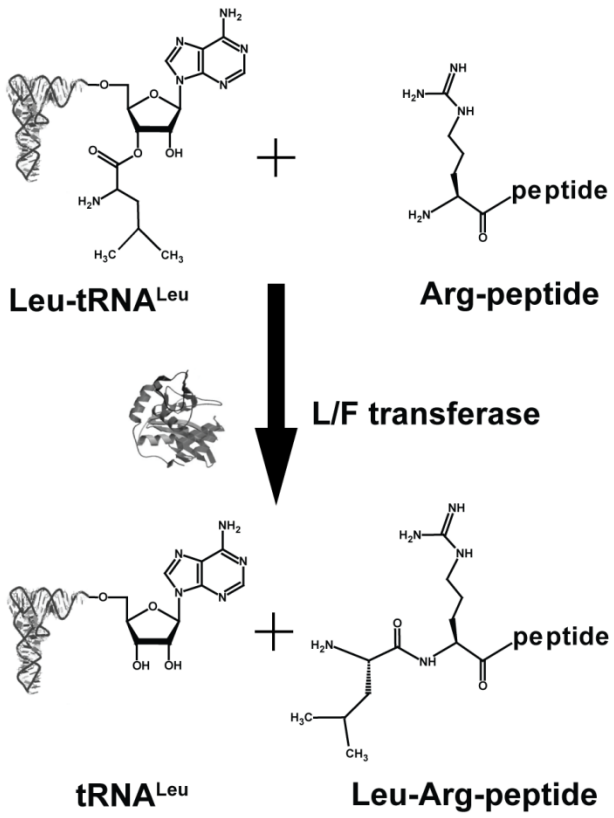

Figure 1-3: A Schematic of tRNA-dependent post-translational addition of amino acids by L/F transferase. L/F transferase catalyzes the addition of the esterified amino acid Leu or Phe or Met from $tRNA^{Leu}$ (shown) or $tRNA^{Phe}$ or $tRNA^{Met}$ to the N-terminal of a protein substrate bearing an Arg (shown) or Lys or Met residue.

(Guo *et al.* 2002). Upon binding to ATP, ClpA oligomerizes and form a ring-shaped hexamer that interacts with the barrel-shaped double-stacked heptameric rings of peptidase ClpP at one or both ends of the barrel (Wang *et al.* 1997, Maglica *et al.* 2009). This oligomeric ClpAP structure is similar to the eukaryotic 26S proteasome. ClpA hydrolyzes ATP and unfolds the substrate protein in a linear fashion into the internal active site chamber of ClpP, which results in the degradation of N-end rule substrates (Sprangers *et al.* 2005).

1.2. The Structure of L/F transferase: the GNAT-like domain and the Dupli-GNAT Superfamily

E. coli L/F transferase is a ~26 kDa monomeric enzyme present in the cytosol. L/F transferase consists of two domains, a small N-terminal domain with a novel fold (residues 1-63: β1, 3_{10}1, β2, α1, β3, β4) and a larger catalytic C-terminal domain (residues 64-234) (**Figure 1-4**) (Dong *et al.* 2007). The larger C-terminal domain is further divided into a conserved core region and a surrounding region. The core region is structurally identical to the GCN5-related *N*-acetyltransferase (GNAT) superfamily fold (β6, α3, α4, antiparallel β7-9, characteristic central α5, β10, α6, β11). The *N*-acetyltransferases from the GNAT superfamily fold catalyzes the transfer of acetyl group from acetyl-CoA to a primary amine, and the reaction is mechanistically similar to the post-translational addition of amino acids (Vetting *et al.* 2005, Rai *et al.* 2006). This conserved core is further surrounded by two α helices (α2, α7), two β strands (β5, β12) and

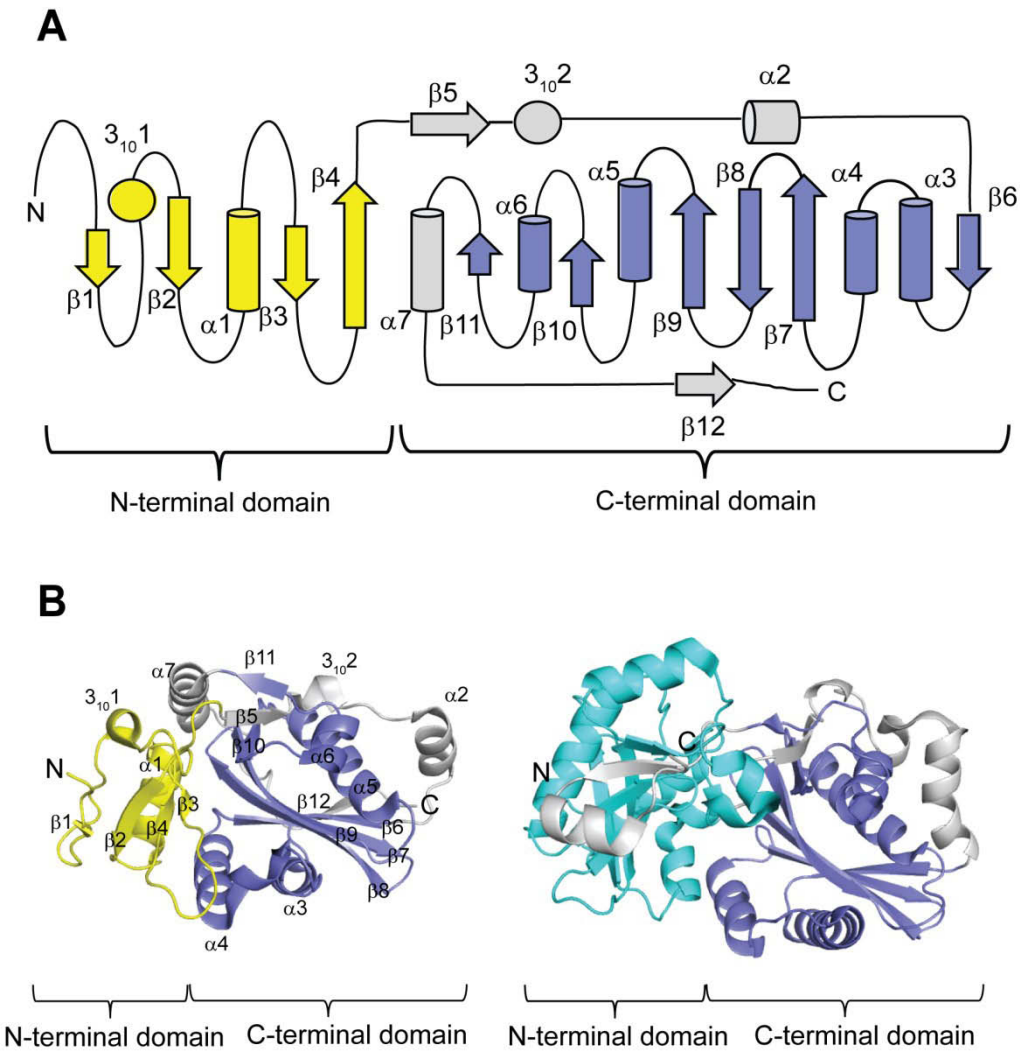

Figure 1-4: The structure of L/F transferase. A) A cartoon representation of the folded topology of L/F transferase, where alpha helices are represented by cylinders, beta sheets as arrows, and 3_{10} helices as circles. The N-terminal domain of L/F transferase (residues 1-63, coloured in yellow), and the C-terminal domain (residues 64-232 coloured in grey with the core GNAT-like domain (residues 87-202) coloured in blue). **B)** A ribbon diagram of (left) L/F transferase (PDB ID: 2Z3K) and (right) $FemX_{Wv}$ (PDB ID: 4II9). L/F transferase contains a single partial GNAT-like domain at the C-terminus, meanwhile $FemX_{Wv}$ contains two complete GNAT-like domains with domain 1 at the N-terminus (coloured in cyan) and domain 2 at the C-terminus (coloured in grey with the core coloured in blue). The figure was generated using PyMOL version 1.41 and is adapted from (Dong *et al.* 2007).

a 3_{10}-helix ($3_{10}2$). Most of the conserved residues reside at the large but shallow central cavity between the interfaces of the two domains. X-ray crystal structures with substrate analogues and product peptide reveal that the central cavity is indeed the site of substrate binding and catalysis (Suto *et al.* 2006, Watanabe *et al.* 2007).

Initially it was believed that there is a lack of sequence similarity between the eubacterial L/F transferase and eukaryotic ATE1 despite similarities in their enzymatic reaction (Balzi *et al.* 1990, Shrader *et al.* 1993, Graciet *et al.* 2006). Through a more sensitive sequence comparison model, ATE1 is demonstrated to be evolutionary related to L/F transferase (Rai *et al.* 2006). L/F transferase and ATE1 are currently grouped together and belong to the "Dupli-GNAT" superfamily that reflects the likely ancestral duplication of the GNAT-like domain, as similarly observed in the *N*-myristoyltransferase family (Weston *et al.* 1998, Rai *et al.* 2006). This "Dupli-GNAT" superfamily of proteins contains either two complete or a single partial version of the GNAT-like domain. Based on modeling, unlike the single GNAT-like domain observed in L/F transferase, ATE1 is predicted to have two complete characteristic GNAT-like domains. The N-terminal domain of ATE1 is predicted to lose its acetyl-CoA co-factor binding function, while the C-terminal domain is predicted to be the catalytic domain and binds an aa-tRNA (Rai *et al.* 2006).

There are other tRNA-dependent transferases, which are not involved in N-end rule-mediated protein degradation, but share the GNAT-

like domain structural homology and catalyze a similar post-translational addition of amino acids enzymatic reaction (**Figure 1-5**). The Fem transferase family (FemABX) are involved in peptidoglycan biosynthesis and they also belong to the "Dupli-GNAT" superfamily (Rai *et al.* 2006). Fem is an abbreviation for "factors essential for methicillin" resistance (Berger-Bachi *et al.* 1989). FemX catalyzes the interpeptide bond formation where an Ala is transferred from alanyl-tRNA to the ε–amino group of Lys_3 of the UDP-MurNAc-pentapeptide (Billot-Klein *et al.* 1997). The Fem transferase family of proteins have two structurally equivalent GNAT-like domains with the presence (ex. FemA of *Staphylococcus aureus* (Benson *et al.* 2002)) or absence (ex. FemX from *Weissella viridescens* (Biarrotte-Sorin *et al.* 2004)) of a coiled-coil domain. The coiled-coil domain has been proposed to be important for tRNA recognition in FemA (Benson *et al.* 2002). Substrate peptide-bound X-ray crystal structure shows that the substrates bind to the cleft on the interface between the two domains, similarly observed in L/F transferase and proposed for ATE1. The acceptor peptide interacts mainly with domain I, while domain II is responsible for tRNA recognition (Biarrotte-Sorin *et al.* 2004). The folding of the two GNAT-like domain of ATE1 is predicted to be similar to the $FemX_{Wv}$ enzyme (**Figure 1-4B**) (Rai *et al.* 2006).

The following tRNA-dependent transferases described below have been shown to contain the GNAT-like domain by structural homology prediction models, despite a lack of sequence homoloy

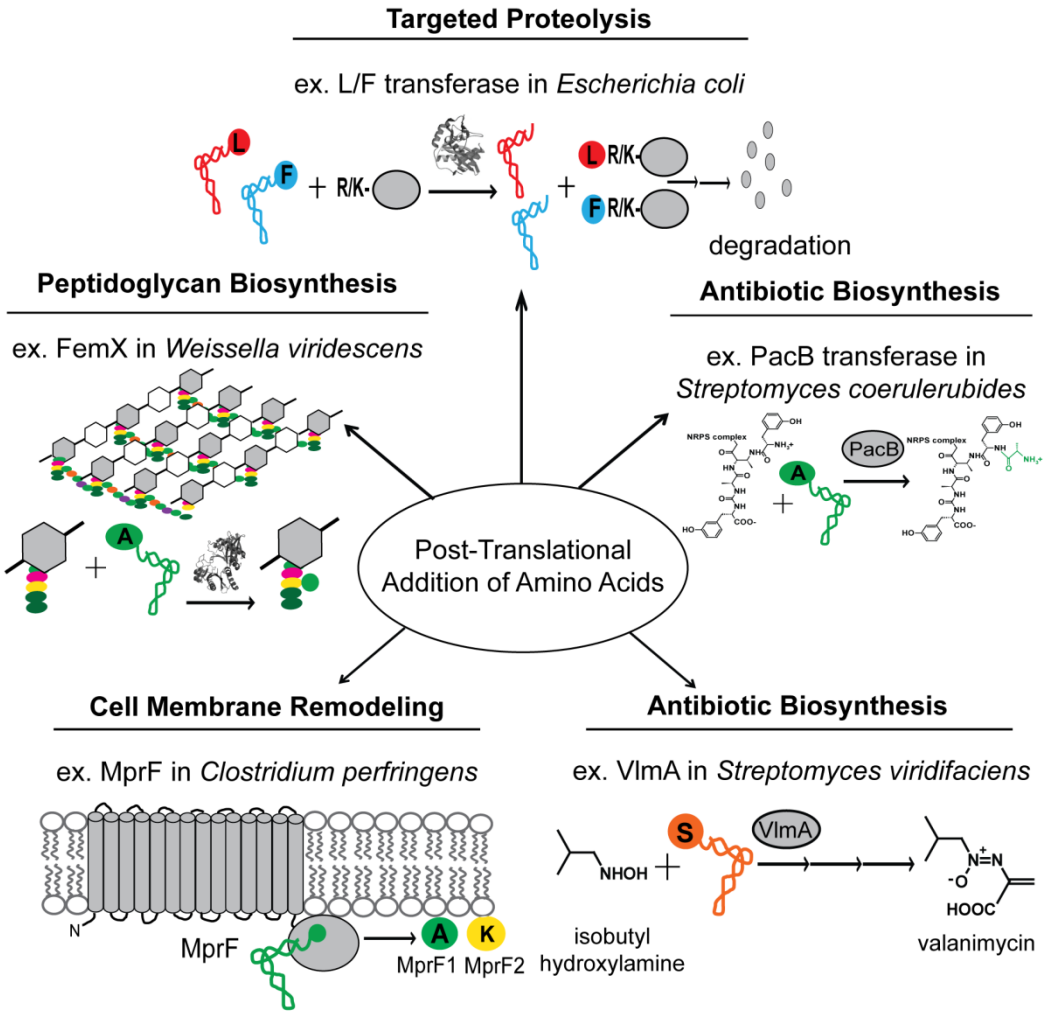

Figure 1-5: The post-translational addition of amino acids by tRNA-dependent transferases with GNAT-like domain. The post-translational addition of amino acid from an aa-tRNA to various substrates by enzymes containing the GNAT-like domain may be a general strategy of biosynthesis and regulation in eubacteria and eukaryotes.

(Iyer *et al.* 2009, Zhang *et al.* 2011). One enzyme is the PacB transferase from *Streptomyces coeruleorubidus* involved in pacidamycin antibiotic biosynthesis, where it adds Ala from Ala-tRNA^{Ala} to the N-terminus of the tetrapeptide intermediate (Zhang *et al.* 2011). PacB transferase is predicted to have two domains separated by a central cleft where the C-terminal resembles the GNAT-like domain important for catalysis (Zhang *et al.* 2011). Pacidamycin has been suggested to inhibit bacterial cell wall assembly (Zhang *et al.* 2011). Another enzyme is the aminoacyl phosphatidylglycerol synthase (aaPGS) encoded by the gene *mprF* (multipeptide resistance factor) (Roy and Ibba 2008, Hebecker *et al.* 2011). aaPGSs are involved in cell membrane permeability remodeling and found in both gram-positive and gram-negative bacteria. MprF1 (or AlaPGS) adds Ala and MprF2 (or LysPGS) adds Lys from cognate aa-tRNA to the polar head group of phosphatidylglycerol (PG) (Roy and Ibba 2008, Hebecker *et al.* 2011). MprF1/2 is predicted to have an N-terminal transmembrane domain with 13 or 14 transmembrane helices and a C-terminal GNAT-like domain for catalysis (Roy and Ibba 2008, Iyer *et al.* 2009, Hebecker *et al.* 2011). VlmA, an MprF homolog from *Streptomyces viridifaciens*, adds Ser from a Ser-tRNA^{Ser} to isobutylhydroxylamine for valanimycin antibiotic biosynthesis (Garg *et al.* 2008). Valanimycin has been suggested to inhibit DNA synthesis in pathogenic bacteria (Yamato *et al.* 1987). Together, the post-translational addition of amino acid from an aa-tRNA to various substrates by enzymes containing the GNAT-like

domain may be a general strategy of biosynthesis and regulation in eubacteria and eukaryotes.

1.3. The Catalytic Mechanism of L/F transferase: tRNA-Dependent Post-Translational Addition of Amino Acids

The non-ribosomal addition of amino acids to proteins was first described almost 50 years ago in *E. coli* (Kaji *et al.* 1965a, Kaji *et al.* 1965b). The tRNA-dependent post-translational addition of amino acids is an acyl transfer reaction, transferring the aminoacyl moiety from the 3' end of an aa-tRNA substrate to an acceptor substrate. This often results in an amide linkage between the donor amino acid and acceptor molecule, where the α-amino group of the N-terminus of a protein polypeptide (ex. L/F transferase and PacB) or the ε-amino group of Lys_3 of a peptidoglycan intermediate (ex. $FemX_{Wv}$) act as nucleophiles and attack the aminoacyl carbonyl group. This process is also described as tRNA-dependent non-ribosomal peptide bond formation. Other examples of addition of amino acids by GNAT-like domain enzymes occurs via ester linkages have also been described, such as the hydroxyl groups of phosphatidylglycerols (ex. AlaPGS) and isobutylhydroxylamine (ex. VlmA) act as nucleophiles (Garg *et al.* 2008, Hebecker *et al.* 2011).

An initial protein-based catalytic mechanism has been proposed for L/F transferase based on available complex structures with substrate analogues and product peptide (Suto *et al.* 2006, Watanabe *et al.* 2007). The proposed mechanism suggests two catalytic residues, D186 and

Q188, on the enzyme that are actively involved in catalyzing the transfer reaction (Watanabe *et al.* 2007). This mechanism is similar to the classic reverse acylation step of proteolysis observed in serine proteases. The general base Q188, first activated by D186 via an electron-relay system, attracts a proton from the α-NH_3^+ group of the N-terminal Arg acceptor substrate peptide and facilitates the nucleophilic attack on the carbonyl carbon of an aa-tRNA donor substrate. Through hydrogen bonding, N191 has been proposed to not only enhance the electrophilicity of the carbonyl carbon of aa-tRNA, but also stabilizes the tetrahedral oxyanion intermediate formed (Watanabe *et al.* 2007). The aminoacyl transfer completes with the protonation of the 3'-oxygen of the deacylated tRNA from Q188, and the two products are released. This proposal of Q188 (amide functional group) acting as a general base is unconventional.

With additional mutagenesis and more sensitive data collection, we have proposed an alternative catalytic mechanism for L/F transferase (see **Chapter 3**) (Fung *et al.* 2011). This alternative catalytic mechanism involves the participation of the aa-tRNA A_{76} 2'-OH where it contributes to catalysis by acting as a general acid/base in proton shuttling, while L/F transferase has a more passive role in the specific binding and positioning of the substrates (Fung *et al.* 2011). The tRNA-dependent non-ribosomal peptide bond formation catalytic mechanism by L/F transferase is more similar to the catalytic mechanism proposed for the ribosome than previously believed.

1.4. The Substrates Specificities of L/F transferase

The only aa-transferase in *E. coli* is L/F transferase which catalyzes the transfer of an Leu or Phe or Met moiety from an cognate aa-tRNA (donor substrate) onto the N-terminus of a protein polypeptide having an N-terminal Arg or Lys or Met residue (acceptor substrate) (Leibowitz and Soffer 1971a, Soffer 1973, Scarpulla *et al.* 1976, Tobias *et al.* 1991, Shrader *et al.* 1993, Ninnis *et al.* 2009). The *aat* (gene encoding L/F transferase) is widely distributed in eubacteria including actinobacteria, proteobacteria, chlorobi, cyanobacteria, spirochaetes, and thermus-deinococcus, despite missing in many species as well (Ichetovkin *et al.* 1997, Varshavsky 2011). *In vitro* characterization of *E. coli* L/F transferase and the most divergent cyanobacterium *Synechocystis sp.* PCC6803 L/F transferase homolog shows that despite significant sequence divergence (80%) between the two enzymes, the donor aa-tRNA and acceptor peptide substrate specificity is highly conserved (Ichetovkin *et al.* 1997).

1.4.1. Acceptor Substrate Specificity: The N-terminus of a Protein Polypeptide

Based on the initial identification of the absolute requirement of a free α-amino group and an N-terminal basic amino acid ($_{L}$-Arg or $_{L}$-Lys) in the acceptor peptide (Soffer 1973), L/F transferase activity assays have largely been based on model *in vitro* substrates, such as bovine serum albumin (BSA) (Leibowitz and Soffer 1970, Rao and Kaji 1974, Deutch

and Soffer 1975), α_{s1}-casein and β-casein (Leibowitz and Soffer 1971a, Abramochkin and Shrader 1996, Ichetovkin *et al.* 1997, Suto *et al.* 2006, Watanabe *et al.* 2007, Ninnis *et al.* 2009, Wagner *et al.* 2011), β-galactosidase (Tobias *et al.* 1991), and synthetic peptides (Soffer 1973, Ebhardt *et al.* 2009, Fung *et al.* 2011, Kawaguchi *et al.* 2013, Fung *et al.* 2014a, Fung *et al.* 2014b).

The availability of the X-ray crystal structure of L/F transferase in complex with the product peptide reveals the molecular recognition of the acceptor peptide (Watanabe *et al.* 2007). The product peptide (α-casein fragment FRYLG, cyan) binds at the central cleft between the N- and C-terminal domains (**Figure 1-6**) (Watanabe *et al.* 2007). There are three deep binding pockets at the cleft, one for the donor amino acid (d_1) and two for the first two residues of the acceptor peptide (a_1 and a_2) (Dougan *et al.* 2010). The donor Phe residue is positioned in the C-shaped hydrophobic pocket (d_1) as described in **1.4.2.** (Watanabe *et al.* 2007). The positively charged Arg residue is bound to the negatively charged pocket (a_1) adjacent to the C-shaped hydrophobic pocket (d_1) (Watanabe *et al.* 2007). The guanidinium group of Arg forms strong electrostatic interaction with E156. Mutagenesis of E156 suggests that E156 dictates the N-terminal basic amino acid specificity of the acceptor peptide (Watanabe *et al.* 2007). Additionally the product peptide forms hydrogen bonding interactions with residues P48, S157, Q188, and N191 (**Figure 1-6C**). Meanwhile W49, W59, and W111 form hydrophobic interactions with

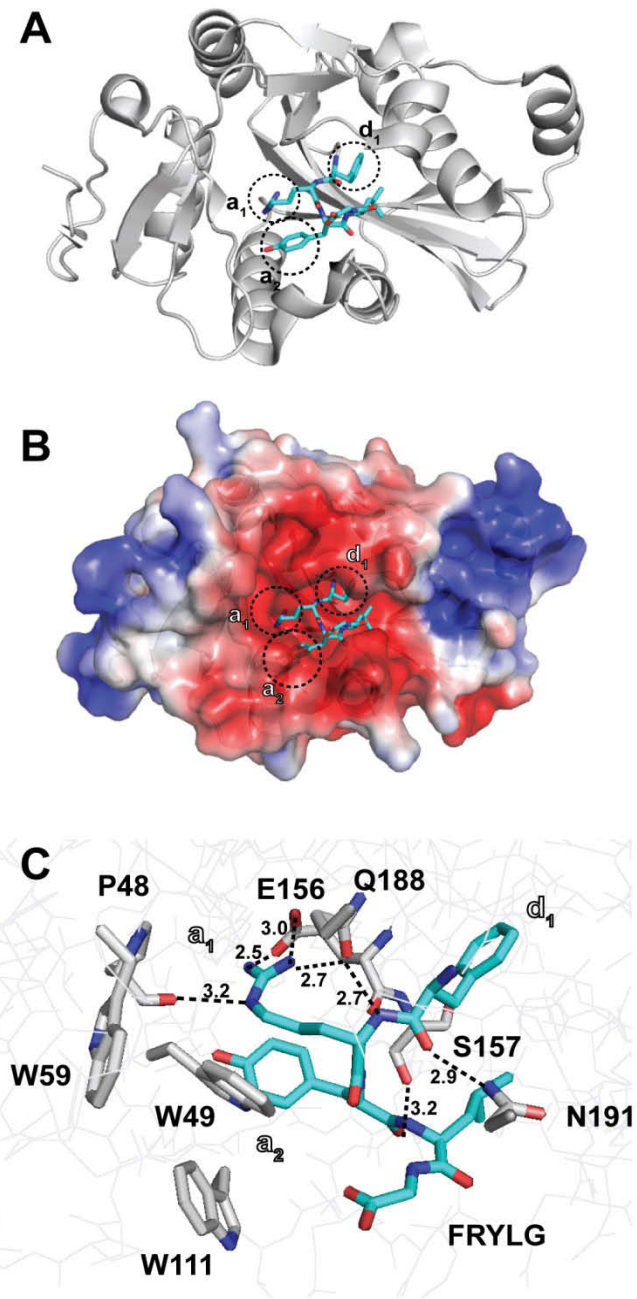

Figure 1-6: Structural basis of peptide substrate recognition by L/F transferase. **A)** Cartoon and **B)** electrostatic surface representation of L/F transferase (PDB ID: 2Z3N) with product peptide FRYLG (cyan, sticks) bound. The three pockets for the donor amino acid (d_1), acceptor peptide position 1 (a_1) and position 2 (a_2) are specified with a dashed line circle. **C)** The key residues interacting with the product peptide are shown as sticks.

Tyr in the a_2 pocket. Based on biochemical data and structural analyses, it is thought that only the N-terminal amino acid residue specificity is essential (a_1), and the remaining of the peptide recognition is recognized in a sequence-independent manner (Watanabe *et al.* 2007). More recently the penultimate specificity (a_2) has been revealed. It has been demonstrated that although the transfer reaction by L/F transferase occurs to all acceptor peptides with N-terminal basic residues, the transfer rates vary depending on the identity of the penultimate residue (Kawaguchi *et al.* 2013). Specifically, small hydrophilic (i.e. Ser, Thr) or basic residues (i.e. Arg, Lys) at the penultimate position (a_2) are more favourable than large hydrophobic (i.e. Trp) or constrained (i.e. Pro) residues (Kawaguchi *et al.* 2013). This marks a more important role for the penultimate residue in acceptor peptide substrate recognition than previously accepted.

L/F transferase acceptor peptide recognition is complementary to the downstream ClpS recognition. The adaptor protein ClpS was suggested to play an important role in determining substrate specificity for the degradation by ClpAP (Dougan *et al.* 2002b, Guo *et al.* 2002, Zeth *et al.* 2002, Erbse *et al.* 2006, Wang *et al.* 2007, Schmidt *et al.* 2009, Schuenemann *et al.* 2009, De Donatis *et al.* 2010). ClpS has a preference of binding to hydrophobic residues including Phe, Leu, Tyr, and Trp, which complements the 1º destabilizing residue identity in the N-end rule (Wang *et al.* 2008a). *In vitro* studies have shown that the L/F transferase product (i.e. FR-protein) is a better ClpS substrate than a protein with only the N-

terminal 1^{o} destabilizing residue (i.e. F-protein) (Erbse *et al.* 2006). The X-ray crystal structure of the cone-shaped ClpS shows a hydrophobic cavity (binds hydrophobic 1º destabilizing residues) lined by negatively charged (D35, D36) and polar residues to hydrogen bond with the first two residues of the N-degron (i.e. F, L, W, Y at d_1 and R or K at a_1) (Guo *et al.* 2002, Wang *et al.* 2008a, Schuenemann *et al.* 2009). ClpS also utilize electrostatic charge in the binding site to modulate acceptor peptide binding through a similar mechanism as L/F transferase (Erbse *et al.* 2006, Wang *et al.* 2008a, Dougan *et al.* 2010). Additionally any basic residue near the N-terminus enhances, while acidic residue decreases, both L/F transferase activity and ClpS recognition (Soffer 1973, Wang *et al.* 2008a, Kawaguchi *et al.* 2013). An unstructured region having a hydrophobic element six to twelve residues downstream of the N-degron is essential for ClpSAP recognition and degradation (**Figure 1-1**) (Wang *et al.* 2008a, Ninnis *et al.* 2009, Schuenemann *et al.* 2009, Dougan *et al.* 2010). Together, the *in vivo* half-life of proteins in *E. coli* depends not only on the N-terminal residue but also on the penultimate residue and an available unstructured hydrophobic element.

The identification of the first *in vivo* substrate of L/F transferase, PATase, unexpectedly did not fit with the current acceptor substrate specificity. PATase was modified at the neutral N-terminal Met residue (stabilizing) instead of the well-characterized basic N-terminal Arg or Lys residue (2º destabilizing) at the a_1 position (Ninnis *et al.* 2009). This

finding cannot be explained by the current electrostatic interaction with E156 model. A hypothesis has been proposed, where the penultimate residue (a_2) of the acceptor peptide may play a role in recognition (Dougan *et al.* 2010). The penultimate residue of PATase is the neutral and hydrophilic Asn, which is a favourable a_2 residue for L/F transferase (Kawaguchi *et al.* 2013). It may be possible for Asn to bind favourably in the a_2 pocket, and compensate for the unfavourable binding of N-terminal Met to the negatively charged a_1 pocket. It would be beneficial to determine whether there are other possible combinations of favourable binding in a_2 that compensate the unfavourable binding in a_1. Specifically the use of peptide libraries with various residue combinations at the first and second (a_1 and a_2) positions would enable this investigation. Interestingly, in the yeast and mouse Arg/N-end rule pathway, the idea that an N-terminal Met is a stabilizing residue has recently been refuted (Kim *et al.* 2014). N-terminal Met is reported to be a *de*stabilizing residue under the condition that it is followed by a bulky, hydrophobic residue (i.e. Leu, Ile, Phe, Tyr, or Trp) in the penultimate position (Kim *et al.* 2014). If a similar conditional degradation of protein substrates bearing an N-terminal Met with penultimate residue of polar or positively charged residues is true for the *E. coli* Leu/N-end rule pathway, it would greatly expand the pool of potential N-end rule protein substrates.

L/F transferase has also been shown to add multiple rounds of Leu to the N-terminus of PATase, followed by the last addition of Phe (Ninnis

et al. 2009, Humbard *et al.* 2013). This suggests that Leu, a 1º destabilizing residue, may also act as a 2º destabilizing residue such that L/F transferase recognizes the hydrophobic Leu in the negatively charged a_1 pocket under certain unknown circumstances. The multiple Leu addition reaction may be terminated by the attachment of Phe to the N-terminus. This suggests that Phe binds unfavourably at the a_1 position. Overall the findings on PATase challenge the current understanding on L/F transferase acceptor peptide substrate specificity, and may imply that L/F transferase has a broader specificity than previously believed.

1.4.2. Donor Substrate Specificity: The Aminoacyl-tRNA

The aa-tRNA substrate utilization by L/F transferase has been demonstrated to include $_L$-Leu-tRNA^{Leu}, $_L$-Phe-tRNA^{Phe}, and to a lesser extent $_L$-Met-tRNA^{Met} *in vitro* (Leibowitz and Soffer 1969, Scarpulla *et al.* 1976, Abramochkin and Shrader 1996). The dominant modification catalyzed by L/F transferase observed *in vivo* is leucylation and a preference for a specific leucine isoacceptor (anticodon 5'-CAG-3') has been suggested (Rao and Kaji 1974, Shrader *et al.* 1993). Despite this specificity for tRNA^{Leu}, the majority of biochemical and structural studies on aa-tRNA specificity have focused on tRNA^{Phe} and its analogues as model substrates (Suto *et al.* 2006, Watanabe *et al.* 2007). Based on the available tRNA substrate analogue-bound X-ray crystal structures and biochemical data, it has been proposed that the aa-tRNA recognition by L/F transferase involves mainly the 3' aminoacyl adenosine (Leibowitz and

Soffer 1971a, Abramochkin and Shrader 1996, Suto *et al.* 2006, Watanabe *et al.* 2007).

The minimal active tRNA substrate analogue phenylalanyl adenosine (rA-Phe) and the inert tRNA substrate analogue puromycin differ by a few chemical differences, including an ester versus amide linkage, a dimethylation modification on the adenine base, and a methoxy modification on the phenylalanine side chain. The X-ray crystal structure of L/F transferase in complex with rA-Phe (green) and puromycin (yellow) show that both analogues bind to the central cleft between the interface of the two domains of L/F transferase (**Figure 1-7A**) (Suto *et al.* 2006, Watanabe *et al.* 2007). At the end of this cleft, a C-shaped hydrophobic pocket (also referred to as the d_1 pocket consisting of M144, F153, L170, F173, I185 residues) interacts with the amino acid side chain of the analogues. The C-shaped hydrophobic pocket dictates its donor amino acid specificity via steric hindrance. Larger β-branched amino acids (i.e. Ile and Val) are sterically hindered from binding, while smaller amino acids (i.e. Ala and Pro) provide insufficient hydrophobic contacts (Suto *et al.* 2006, Watanabe *et al.* 2007).

With regards to the regiospecificity of aa-tRNA recognition, L/F transferase only utilizes 3'-aa-tRNAs, as 2'-Phe-$tRNA^{Phe}$ (with 3' deoxyadenosine) was demonstrated to not be a substrate for L/F transferase (Watanabe *et al.* 2007). X-ray crystal structures reveal that both rA-Phe and puromycin adopt a C_3' endo conformation when bound to

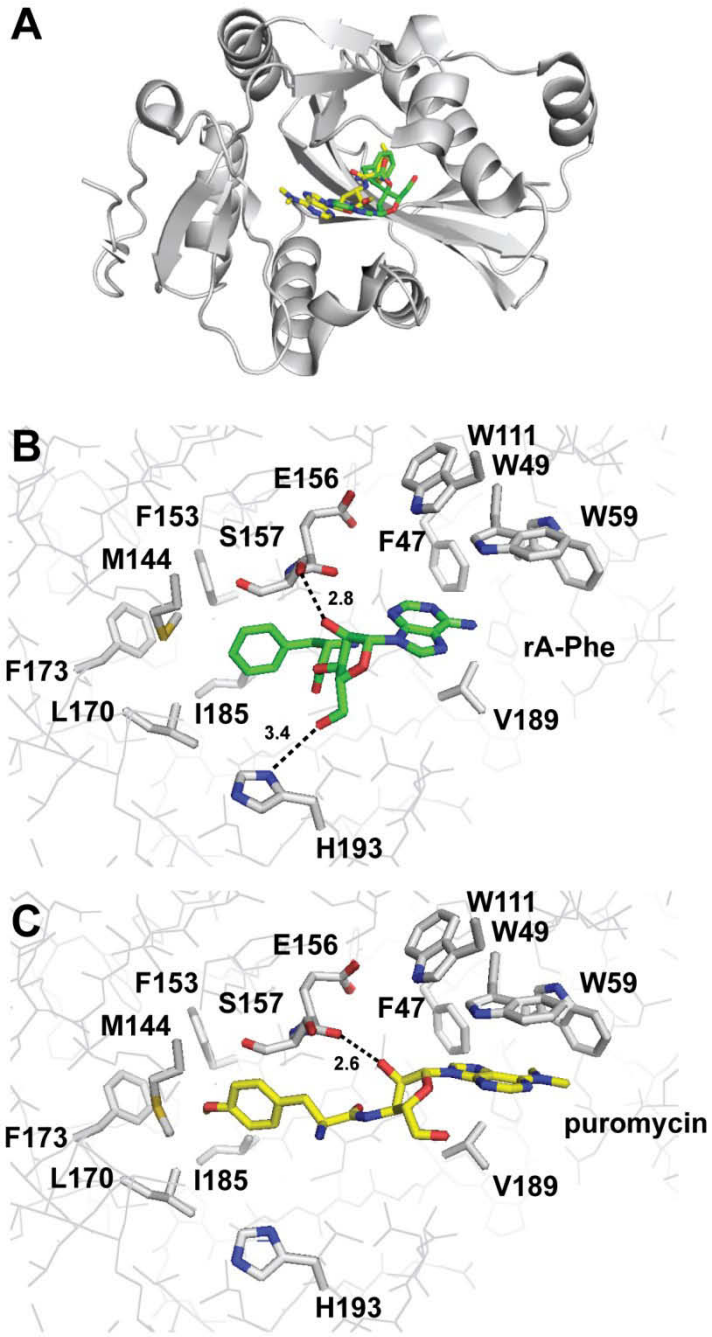

Figure 1-7: Structural basis of tRNA substrate recognition by L/F transferase. **A)** Cartoon representation of L/F transferase (PDB ID: 2Z3K) with tRNA substrate analogues rA-Phe (green, sticks, PDB ID: 2Z3K) and puromycin (yellow, sticks, PDB ID: 2DPT) bound. The key residues interacting with the tRNA substrate analogues are shown as sticks for **B)** rA-Phe-bound and **C)** puromycin-bound.

L/F transferase, which is the conventional conformation of a ribose of an RNA (Suto *et al.* 2006, Watanabe *et al.* 2007). The adenine base is stabilized by π-π stacking interaction with W49, which is further positioned by interactions with W111 (Suto *et al.* 2006, Watanabe *et al.* 2007). In the puromycin structure, additional hydrophobic contacts between L/F transferase (F47, W59, and V189) and 6-*N*,*N*-dimethyladenosine further stabilize binding (Suto *et al.* 2006).

Although the amino acid side chain of both rA-Phe and puromycin binds to the amino acid binding C-shaped pocket, closer examination of the structures with the two analogues revealed differences in the binding conformations of the ribose and adenine base (see **Chapter 4**) (**Figure 1-7B** and **C**) (Fung *et al.* 2014a). Through mutagenesis and assay with an additional tRNA substrate analogue (synthesis in collaboration with Dr. Kollappillil Krishnakumar and Dr. Peter Strazewski at the Université de Lyon), we have determined that both rA-Phe-amide and puromycin bind to L/F transferase with the same order of affinity, despite differences in their binding conformation. This is due to the different modifications on puromycin having opposite effects on its interaction with L/F transferase (Fung *et al.* 2014a). Additionally, we were able to illustrate that mutations to M144 that enlarge the C-shaped pocket enhance puromycin binding (Fung *et al.* 2014a). This study will aid in the future design of stronger binding substrate analogues for studying the molecular mechanisms of L/F transferase.

The study of aa-tRNA recognition by L/F transferase has been limited to the 3' end, since it has been shown that L/F transferase shows minimal specificity to the tRNA body. Misacylated-tRNAs (such as Met-tRNA^{Val} and Phe-tRNA^{Val}), oligo-DNA-RNA hybrids, and minimal substrate analogues aminoacyl-ribonucleoadenosine (rA-aa) all act as substrates for L/F transferase (Leibowitz and Soffer 1971a, Abramochkin and Shrader 1996, Watanabe *et al.* 2007, Wagner *et al.* 2011). Additionally, mutations to the anticodon and variable loop of an aa-tRNA do not result in measurable effects on enzyme activity (Leibowitz and Soffer 1971a, Abramochkin and Shrader 1996). However, comparison of the reported K_Ms of rA-Phe (124 μM) and Phe-tRNA^{Phe} (2 μM) suggests that the tRNA body contributes significantly to binding (Rao and Kaji 1974, Abramochkin and Shrader 1996, Wagner *et al.* 2011).

Through *in vitro* transcribed tRNA hybrids and kinetic assays, we have investigated the molecular basis of the preference for a specific tRNA^{Leu} isoacceptor. We determined that L/F transferase specifically recognizes two independent sequence elements in the acceptor stem of an aa-tRNA in addition to the 3' rA-aa (see **Chapter 5**) (Fung *et al.* 2014b). This study illustrates that the tRNA body contributes to aa-tRNA substrate recognition significantly in a sequence-dependent manner. Although an X-ray crystal structure of an intact aa-tRNA in complex with L/F transferase was not solved, we fine-tuned and proposed a new aa-tRNA recognition model where the positively-charged cluster of residues

(R76, R80, K83, R84 on α2 helix) modulates aa-tRNA specificity through recognition of the acceptor stem (Suto *et al.* 2006). This model may also be extended to other tRNA-dependent GNAT-like domain containing enzymes.

1.4.3. Other Aminoacyl-tRNA Protein Transferases in N-end Rule Pathway

Recall that in eukaryotes ATE1 catalyzes the transfer of Arg from $_{L}$-Arg-tRNA^{Arg} onto the N-terminus of proteins having an Asp, Glu, or modified Cys (Gonda *et al.* 1989). Mutagenesis directed by the model structure, Rai *et al*. identified a universally conserved Lys at position 417, in human ATE1 that is important for acceptor peptide substrate specificity via an analogous electrostatic interaction as described for L/F transferase for the recognition of peptides with basic N-termini (Rai *et al.* 2006). Additionally, there are also three cysteine residues (C20, C23, C94/95) near the N-terminus of yeast ATE1 that have been shown to be important for activity (Li and Pickart 1995), as well as its ability to induce degradation in yeast (Kwon *et al.* 1999).

Additionally as a result of alternative splicing, mammals express multiple forms of ATE1 which results in different substrate specificity, tissue expression, and localization. For example, while mouse ATE1-1 and ATE1-2 can arginylate N-terminal Asp, Glu, and oxidized Cys, mouse ATE1-3 and ATE1-4 specifically arginylate N-terminal oxidized Cys only (Kwon *et al.* 1999, Rai and Kashina 2005). Cys-specific arginylation

occurs in mammals but not in yeast maybe due to the requirement of Cys to be modified to cysteic acid prior to arginylation (Kwon *et al.* 2002). A set of three amino acids (F15, E16, G17) have been suggested to dictate the substrate specificity of the different mouse ATE1 forms (present in ATE1-1/2 but absent in ATE1-3/4) (Rai *et al.* 2006). Despite these reports, there is no catalytic mechanism or detailed substrate specificities reported for ATE1. Systematic kinetic and mechanistic characterization of ATE1 will further expand the molecular insights for this enzyme.

Other aa-transferases involved in the N-end rule pathway with hybrid specificities have also been identified. The prokaryotic leucyl transferase Bpt (bacterial protein transferase, $L^{D,E}$ transferase) is identified in the human pathogen *Vibrio vulnificus*, a gram-negative proteobacterium found in shellfish and may cause gastrointestinal and bloodstream infections in humans (Graciet *et al.* 2006). The ~27 kDa Bpt protein is a 'sequelog' (sequence similarity without implications on evolution) of ATE1 and exhibits hybrid specificity, where it transfers a Leu (prokaryotic donor specificity) to the N-terminal Asp or Glu residue (eukaryotic acceptor specificity) (Varshavsky 2004, Graciet *et al.* 2006). Interestingly, L/F transferase and Bpt are encoded by the *aat-bpt* operon in *V. vulnificus*, while only L/F transferase is encoded in *E. coli*. This suggests that the N-end rule acceptor protein specificity is expanded in *V. vulnificus*. A set of conserved tyrosine residues (Y58, Y170, Y205) and cysteine residues (C18, C65), which are also similarly conserved in ATE1,

seems to be important for Bpt activity (Graciet *et al.* 2006, Varshavsky 2011). Another aa-transferase with hybrid specificity has been identified in the eukaryotic pathogen *Plasmodium falciparum*, a protozoan parasite that may cause malaria. The ~43 kDa ATEL1 (ATE-like, $R^{D, E}$ transferase) is a 'sequelog' of L/F transferase but has identical substrate specificities as ATE1 (Graciet *et al.* 2006). Currently, there is a lack of molecular details with regards to these hybrid tRNA-dependent N-end rule aa-transferases.

Archaea and gram-positive bacteria appear to lack the Leu/N-end rule pathway as they lack sequelogs of *aat*, *bpt*, and *clpS* (Ichetovkin *et al.* 1997, Varshavsky 2011). Phylogenetic analyses of the 'sequelogous' prokaryotic/ eukaryotic L/F transferase/ATEL1 pair and prokaryotic/ eukaryotic Bpt/ATE1 pair suggests that aa-transferases arose twice independently through convergent evolution (Graciet *et al.* 2006). Additionally, the confinement of Leu-conjugation in prokaryotes (L/F transferase and Bpt) and Arg-conjugation in eukaryotes (ATEL1 and ATE1) suggests that Arg as a 1º destabilizing residue occurs later in the evolution of the N-end rule pathway (Graciet *et al.* 2006). Overall, these newly identified aa-transferases substrate specificities expand the potential substrate pool and physiological functions in a variety of species, suggesting that the N-end rule is an evolutionary conserved pathway. We hypothesize that these hybrid aa-transferases would maintain the GNAT-like domain. The study of the catalytic mechanisms and substrate

specificities amongst L/F transferase and ATE1 may be useful for targeting these pathogens.

1.5. The *in vivo* Substrates and Physiological Functions of L/F transferase and *E. coli* N-end Rule

Despite the wide variety of physiological functions identified for the eukaryotic ATE1 and N-end rule (see **1.1.2.**), the physiological functions of the prokaryotic L/F transferase and N-end rule remain enigmatic. Preliminary physiological functions for the *E. coli* N-end rule have been suggested. The *aat* gene resides at the end of a three gene operon with *cydC* and *cydD* (encoding cysteine ABC transporter ATPase responsible for glutathione and cysteine export), which are transcriptionally coupled and *cydCD* have been demonstrated to be essential for *E. coli* cells to exit stationary phase (Shrader *et al.* 1993, Siegele and Kolter 1993, Abramochkin and Shrader 1995). Although the *aat* gene and *clpA*-*clpS* genes are located less than 1 kb apart, they are oriented in opposite direction and are not transcriptionally coupled (Shrader *et al.* 1993). Based on the phenotypes of Δ*aat* and Δ*clpP* cells, *E. coli* N-end rule was proposed to have putative functions in proline catabolism, peptide transport, and growth phase regulation (Soffer and Savage 1974, Deutch and Soffer 1975, Deutch *et al.* 1977, Weichart *et al.* 2003). This initial proposal of the physiological functions is inadequate as the deletion mutations in the strains investigated also had multiple genes deleted in the chromosomal region.

The elusive physiological function of L/F transferase may be revealed by identifying interacting proteins of L/F transferase. Two recent high throughput investigations (Ni^{2+}-NTA pull down and yeast two-hybrid screen) mapping the *E. coli* protein-protein interactome have used L/F transferase as a bait protein (**Table 1-1**) (Arifuzzaman *et al.* 2006, Rajagopala *et al.* 2014). Many putative interacting partners of L/F transferase belong to protein complexes and have wide physiological functions including DNA replication, translation, and metabolism. However these putative interacting proteins of L/F transferase have not been validated and their biological significance as interaction partners remain unclear.

The physiological functions of L/F transferase and N-end rule in *E. coli* remain enigmatic partly due to the lack of characterization of its *in vivo* substrates. Preliminary attempts in identifying putative L/F transferase substrates found 21 soluble and 3 ribosomal acceptor proteins via the acylation of Δ*aat* lysates with [^{14}C]-phenylalanine after the addition of purified L/F transferase (Leibowitz and Soffer 1971b, Soffer and Savage 1974). Since that study, most L/F transferase studies focused on model substrates and mechanisms.

The ATP-dependent proteasome-like ClpAP complex has been shown to recognize and degrade model N-end rule substrates, however endogenous substrates that are primarily targeted for ClpAP remain unknown (Tobias *et al.* 1991, Ninnis *et al.* 2009). The identification of

Table 1-1: List of putative proteins that interact with L/F transferase.

Uniprot Accession Number	Gene	Protein Name	Complex	Physiological Function
Large-scale pull down His_6-tagged proteins (bait: His_6-L/F transferase) with Ni^{2+}-NTA columns. Proteins were identified by MALDI-TOF MS. (Arifuzzaman *et al.* 2006)				
P69222	infA	Translation initiation factor IF-1		Translation
P03960	kdpB	Potassium-transporting ATPase B chain	Multimeric	Ion Transport
Q47270	ninE	Protein NinE homolog from lambdoid prophage DLP12		Phage
P21170	speA	Biosynthetic arginine decarboxylase	Homo-tetramer	Amino Acid Biosynthesis
P0DM85	yjdA	Clamp-binding protein CrfC	Homo-oligomers	DNA Replication
P33363 [a]	bglX	Periplasmic beta-glucosidase		Hydrolysis
P0A6F5 [a]	groL	GroEL	Multimeric	Chaperone
Yeast two hybrid binary screen (bait: L/F transferase) of protein-protein interactions in *E. coli* (Rajagopala *et al.* 2014)				
P0AA37	rluA	Ribosomal large subunit pseudouridine synthase A		rRNA processing
P0A8P3 [b]	yggX	Probable Fe(2+)-trafficking protein	Monomer	Oxidative Stress

[a] Denotes identification in control experiment as well

[b] Identified when bait is L/F transferase, but not reversed.

protease substrates has been difficult due to their low cellular concentrations and as they are often targeted for degradation by several proteases *in vivo* (Gur *et al.* 2011). With the identification that the adaptor protein ClpS modulates the substrate specificity for the ClpAP protease complex (Erbse *et al.* 2006), it has been rationalized that proteins that interact with ClpS are N-end rule substrates. Two studies have identified ClpS-interacting proteins during exponential growth at 30 ºC or stationary growth (26 hours) at 37 ºC (Ninnis *et al.* 2009, Schmidt *et al.* 2009). Of the twenty or so identified ClpS-interacting substrates, there are three common proteins identified: DNA protection during starvation (Dps), putrescine aminotransferase (PATase), and proline utilization protein A (PutA) (Ninnis *et al.* 2009, Schmidt *et al.* 2009). Only Dps and PATase have been verified as *bona fide E. coli* N-end rule substrates.

Dps exists *in vivo* as a homomeric dodecamer and protects DNA during starvation and oxidative stress by forming a complex with DNA or by chelating iron from toxic by-products of the Fenton reaction (Wolf *et al.* 1999). In bacteria, the Dps levels are low during exponential growth but increases upon starvation and oxidative stress (Ali Azam *et al.* 1999). Dps proteolysis is rapidly resumed upon nutrient upshift from stationary phase (Stephani *et al.* 2003). There are two alternative forms of Dps, the full length uncleaved $Dps_{2\text{-}167}$ and cleaved $Dps_{6\text{-}167}$. It has been shown that $Dps_{2\text{-}167}$ with an N-terminal Ser is degraded by a different AAA+ protease ClpXP during exponential growth and stationary phase exit (Flynn *et al.*

2003, Stephani *et al.* 2003) . This growth-phase dependent degradation of Dps provides a rapid mechanism in the regulation of Dps activity. Meanwhile Dps_{6-167} with an N-terminal Leu (1º destabilizing) is recognized by ClpS directly and degraded by ClpAP (Ninnis *et al.* 2009, Schmidt *et al.* 2009). The removal of the first five N-terminal amino acids by an unknown endopeptidase abolishes DNA binding and degradation by ClpXP, but maintains ClpSAP-mediated degradation (Flynn *et al.* 2003, Ninnis *et al.* 2009, Schmidt *et al.* 2009). The physiological function of Dps degradation by the interplay of ClpXP and ClpSAP remains to be understood. The targeting of two N-terminal degradation signals in close proximity may ensure the degradation of Dps under different environmental signals.

PATase catalyzes the aminotransferase reaction to convert putrescine to 2-oxoglutarate to generate $_{L}$-glutamate and 4-aminobutanal (Samsonova *et al.* 2003). Putrescine is a polyamine and polyamines have been shown to be involved in protein biosynthesis, oxidative stress and biofilm formation (Tabor and Tabor 1985, Igarashi and Kashiwagi 2000, Chattopadhyay *et al.* 2003, Wortham *et al.* 2007). PATase is the first *in vivo* substrate whose protein levels have been shown to depend on both L/F transferase and the ClpSAP machinery (Ninnis *et al.* 2009, Schmidt *et al.* 2009). The degradation of PATase via the N-end rule may be a mechanism to ensure putrescine homeostasis. However this substrate did not coincide with the substrate specificity initially defined (see **1.4.1.**). The unconventional modification may not represent the general method for the

generation of degradation signal *in vivo*. This unexpected finding calls for a more vigorous study of L/F transferase substrate characterization.

PutA is a multifunctional protein that exists *in vivo* as a homodimer. Not only does it oxidize proline to glutamate for use as carbon and nitrogen source when associated to the membrane, it also function as a transcriptional repressor of the *put* operon (including *putA* and *putP*, sodium-proline transporter) when in the cytosol (Scarpulla and Soffer 1978, Brown and Wood 1992, Zhou *et al.* 2008). The fragment of PutA that is interacting with ClpS is equivalent to the size of PutA lacking the N-terminal DNA-binding domain (Ninnis *et al.* 2009). This suggests that PutA is also cleaved by an unknown endopeptidase, which permits the degradation of the C-terminal catalytic domain by the N-end rule pathway. Interestingly, the Δ*aat* mutant strain also shows increased proline catabolism suggesting that N-end rule may play a role in proline catabolism (Deutch and Soffer 1975, Scarpulla and Soffer 1979).

Recently, a more thorough study identified over 100 putative *E. coli* N-end rule substrates using ClpS affinity column under non-denaturing conditions during mid-logarithmic (O.D. 0.7) and stationary phase growth (16 hours) at 37 ºC (Humbard *et al.* 2013). Edman degradation confirms a majority (32 of 37 tested) of the ClpS-interacting proteins eluted contain 1º destabilizing N-degrons. Approximately 25% of ClpS-interacting proteins are modified in the cell by L/F transferase, which is consistent with a previous study (Soffer and Savage 1974, Humbard *et al.* 2013). The

authors divided the identified putative N-end rule substrate proteins into L/F transferase-independent and -dependent and proposed two potential protease cleavage motifs (see **1.1.1.** and **Figure 1-2B**). ClpS-interacting L/F transferase-independent substrates include ribosomal structural proteins (S1, S2, L1, L4, L2, L7/L12, L15, and L21), RNA polymerase β and β' subunits and more (Humbard *et al.* 2013). L/F transferase-dependent substrates of ClpS include acetyl-CoA carboxylase AccD, ATP synthase AtpA/AtpD, elongation factor EF-Tu, initiator factor IF-2, DNA gyrase GyrB, putative ribosome biogenesis GTPase RsgA, 2-oxoglutarate decarboxylase SucA, trigger factor Tig, and nucleoid-associated protein YbaB (Humbard *et al.* 2013).

Despite differences in the immunoprecipitation methods, growth conditions and time of growth between the three ClpS-interacting protein identification studies (**Table 1-2**), a number of proteins including Dps, PATase, and PutA were identified in more than one study (**Figure 1-8**) (Ninnis *et al.* 2009, Schmidt *et al.* 2009, Humbard *et al.* 2013). Future experiments remain to validate these identified putative substrates as *bona fide E. coli* N-end rule substrates. Many of the identified putative L/F-dependent substrates belong to large protein complexes such as acetyl-CoA carboxylase complex AccA/AccD, ATP synthase AtpA/AtpD, Dps, and 2-oxoglutarate dehydrogenase SucA (Humbard *et al.* 2013). *E. coli* N-end rule may play a role in the remodelling or quality control of protein complexes. Additionally certain N-end rule substrates are

Table 1-2: A comparison of the ClpS-interacting protein identification studies.

Cell lysates	Growth Conditions	Media	Bait	Beads	Elute	MS/MS method	Ref
Δ*clpS*	mid-logarithmic growth at 30 ºC	LB	GST-ClpS and GST-$ClpS_{DD/AA}$	GST sepharose	FR or MS dipeptide	• 1D-in-gel LC-MS/MS	(Schmidt *et al.* 2009)
WT, Δ*clpA*, Δ*aat*	stationary phase growth at 37 ºC (26 hrs)	LB	ClpS-His_6 and $ClpS_{DD/AA}$-His_6	Ni-NTA agarose	FR or MS dipeptide	• 2D-DiGE LC-MS/MS • Mascot search • HCT^{ultra} ion trap	(Ninnis *et al.* 2009)
WT, Δ*clpSA*, Δ*aat*	mid-logarithmic growth at 37 ºC (0.7 O.D.) and stationary phase growth at 37 ºC (16 hrs)	LB	ClpS and $ClpS_{DD/AA}$	AminoLink Plus agarose	FKTA tetrapeptide	• 1D-in-gel LC-MS/MS • Mascot search • LTQ ion trap	(Humbard *et al.* 2013)

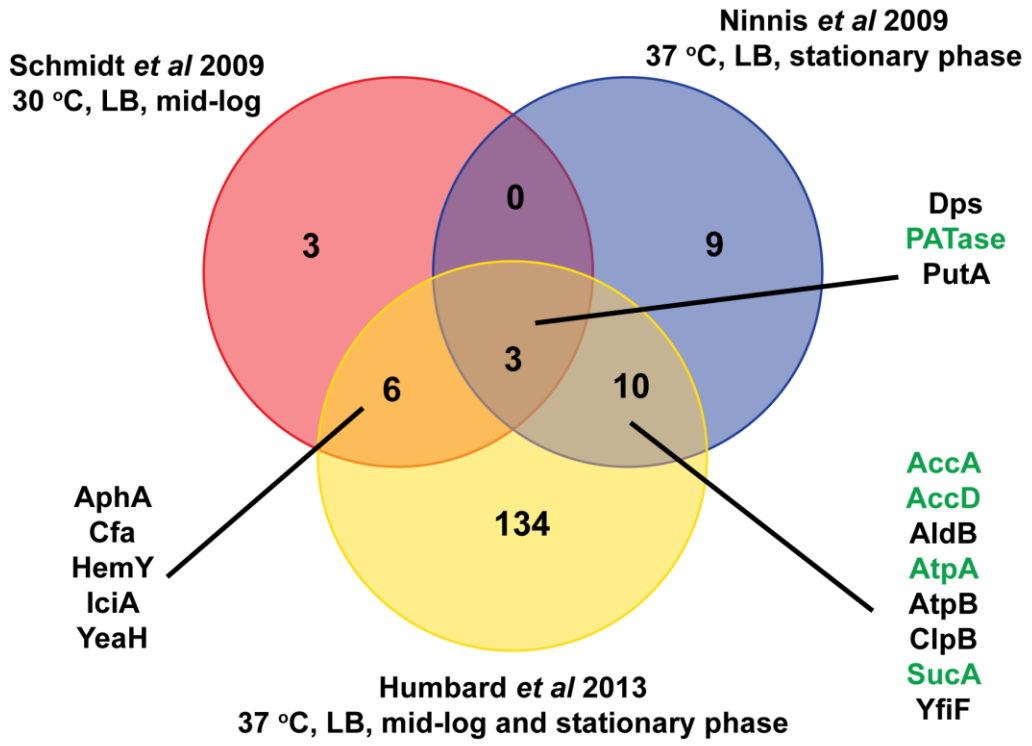

Figure 1-8: Venn diagram comparing the identification of putative *E. coli* N-end rule substrates between three studies. Red circle represents data from (Schmidt *et al.* 2009), blue circle represents data from (Ninnis *et al.* 2009), and yellow circle represents data from (Humbard *et al.* 2013). Some substrates identified have been listed, where L/F-independent substrates are in black and L/F-dependent substrates are in green.

observed to be enriched during exponential growth, while others are enriched during stationary phase. For example, Cfa, SucA and PutA are enriched in exponential growth, meanwhile Dps, AtpA, PATase are enriched during stationary phase. However there is no correlation between LF transferase dependence to the different growth phases. This suggests that N-end rule is a widespread and general mechanism of proteolytic processing under different growth conditions in *E. coli*. Therefore, *E. coli* N-end rule plays a more central role in biological processes than previously recognized including cell division, DNA replication, transcription, translation, metabolism, and protein quality control (Humbard *et al.* 2013).

1.6. Research Goals and Thesis Organization

The goals of this thesis are to elucidate the catalytic mechanism, substrate analogue design, and tRNA substrate recognition of the tRNA-dependent post-translational addition of amino acids catalyzed by the *E. coli* N-end rule L/F transferase. **Chapter 2** summarizes the materials and methods used throughout the thesis. **Chapter 3** presents the investigation on the functional role of the conserved D186 within the active site, where D186 does not participating in catalysis but aids in binding and orienting substrates. This allows the proposal of an alternative 'substrate-assisted' proton shuttling catalytic mechanism. **Chapter 4** describes the study to probe the differential binding of tRNA substrate analogues to L/F transferase as observed in the X-ray crystal structures. The findings

provide insights into the design and development of improved substrate analogues. **Chapter 5** reports the identification and characterization of recognition nucleotides in the acceptor stem of an aa-tRNA substrate. This sheds light on the critical importance of the tRNA body, in addition to the 3' aminoacyl adenosine, in tRNA recognition by L/F transferase. **Chapter 6** summarizes the results from our studies and provides perspectives into the molecular insights and future directions. Finally, **Appendices 1** and **2** describe the work done in the identification of L/F transferase interacting proteins and *in vivo* substrates through affinity purification coupled- and click chemistry coupled- mass spectrometry methods, respectively. The recent findings on L/F transferase catalytic mechanism, substrate specificities and *in vivo* substrates allow hypotheses into the yet enigmatic molecular mechanisms in the initiation of eubacterial N-end rule, the proteases/peptidases responsible for proteolytic cleavage, and physiological functions.

1.7. References

Aasland, R., Abrams, C., Ampe, C., Ball, L.J., Bedford, M.T., Cesareni, G., Gimona, M., Hurley, J.H., Jarchau, T., Lehto, V.P., Lemmon, M.A., Linding, R., Mayer, B.J., Nagai, M., Sudol, M., Walter, U., and Winder, S.J. (2002) Normalization of nomenclature for peptide motifs as ligands of modular protein domains. *FEBS Lett.* **513,** 141-144

Abramochkin, G., and Shrader, T.E. (1995) The leucyl/phenylalanyl-tRNA-protein transferase. Overexpression and characterization of substrate recognition, domain structure, and secondary structure. *J.Biol.Chem.* **270,** 20621-20628

Abramochkin, G., and Shrader, T.E. (1996) Aminoacyl-tRNA recognition by the leucyl/phenylalanyl-tRNA-protein transferase. *J.Biol.Chem.* **271,** 22901-22907

Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S., and Ishihama, A. (1999) Growth phase-dependent variation in protein composition of the *Escherichia coli* nucleoid. *J.Bacteriol.* **181,** 6361-6370

Arifuzzaman, M., Maeda, M., Itoh, A., Nishikata, K., Takita, C., Saito, R., Ara, T., Nakahigashi, K., Huang, H.C., Hirai, A., Tsuzuki, K., Nakamura, S., Altaf-Ul-Amin, M., Oshima, T., Baba, T., Yamamoto, N., Kawamura, T., Ioka-Nakamichi, T., Kitagawa, M., Tomita, M., Kanaya, S., Wada, C., and Mori, H. (2006) Large-scale identification of protein-protein interaction of *Escherichia coli* K-12. *Genome Res.* **16,** 686-691

Bachmair, A., Finley, D., and Varshavsky, A. (1986) In vivo half-life of a protein is a function of its amino-terminal residue. *Science.* **234,** 179-186

Bachmair, A., and Varshavsky, A. (1989) The degradation signal in a short-lived protein. *Cell.* **56,** 1019-1032

Balzi, E., Choder, M., Chen, W.N., Varshavsky, A., and Goffeau, A. (1990) Cloning and functional analysis of the arginyl-tRNA-protein transferase gene ATE1 of *Saccharomyces cerevisiae. J.Biol.Chem.* **265,** 7464-7471

Benson, T.E., Prince, D.B., Mutchler, V.T., Curry, K.A., Ho, A.M., Sarver, R.W., Hagadorn, J.C., Choi, G.H., and Garlick, R.L. (2002) X-ray crystal structure of *Staphylococcus aureus* FemA. *Structure.* **10,** 1107-1115

Berger-Bachi, B., Barberis-Maino, L., Strassle, A., and Kayser, F.H. (1989) FemA, a host-mediated factor essential for methicillin resistance in Staphylococcus aureus: molecular cloning and characterization. *Mol.Gen.Genet.* **219,** 263-269

Biarrotte-Sorin, S., Maillard, A.P., Delettre, J., Sougakoff, W., Arthur, M., and Mayer, C. (2004) Crystal structures of Weissella viridescens FemX and its complex with UDP-MurNAc-pentapeptide: insights into FemABX family substrates recognition. *Structure.* **12,** 257-267

Billot-Klein, D., Shlaes, D., Bryant, D., Bell, D., Legrand, R., Gutmann, L., and van Heijenoort, J. (1997) Presence of UDP-N-acetylmuramyl-hexapeptides and -heptapeptides in enterococci and staphylococci after treatment with ramoplanin, tunicamycin, or vancomycin. *J.Bacteriol.* **179,** 4684-4688

Bongiovanni, G., Fidelio, G.D., Barra, H.S., and Hallak, M.E. (1995) The post-translational incorporation of arginine into a beta-amyloid peptide increases the probability of alpha-helix formation. *Neuroreport.* **7,** 326-328

Brown, E.D., and Wood, J.M. (1992) Redesigned purification yields a fully functional PutA protein dimer from *Escherichia coli. J.Biol.Chem.* **267,** 13086-13092

Butler, S.M., Festa, R.A., Pearce, M.J., and Darwin, K.H. (2006) Self-compartmentalized bacterial proteases and pathogenesis. *Mol.Microbiol.* **60,** 553-562

Chattopadhyay, M.K., Tabor, C.W., and Tabor, H. (2003) Polyamines protect *Escherichia coli* cells from the toxic effect of oxygen. *Proc.Natl.Acad.Sci.U.S.A.* **100,** 2261-2265

De Donatis, G.M., Singh, S.K., Viswanathan, S., and Maurizi, M.R. (2010) A single ClpS monomer is sufficient to direct the activity of the ClpA hexamer. *J.Biol.Chem.* **285,** 8771-8781

Deutch, C.E., Scarpulla, R.C., Sonnenblick, E.B., and Soffer, R.L. (1977) Pleiotropic phenotype of an *Escherichia coli* mutant lacking leucyl-, phenylalanyl-transfer ribonucleic acid-protein transferase. *J.Bacteriol.* **129,** 544-546

Deutch, C.E., and Soffer, R.L. (1975) Regulation of proline catabolism by leucyl,phenylalanyl-tRNA-protein transferase. *Proc.Natl.Acad.Sci.U.S.A.* **72,** 405-408

Ditzel, M., Wilson, R., Tenev, T., Zachariou, A., Paul, A., Deas, E., and Meier, P. (2003) Degradation of DIAP1 by the N-end rule pathway is essential for regulating apoptosis. *Nat.Cell Biol.* **5,** 467-473

Dong, X., Kato-Murayama, M., Muramatsu, T., Mori, H., Shirouzu, M., Bessho, Y., and Yokoyama, S. (2007) The crystal structure of leucyl/phenylalanyl-tRNA-protein transferase from *Escherichia coli. Protein Sci.* **16,** 528-534

Dougan, D.A., Micevski, D., and Truscott, K.N. (2012) The N-end rule pathway: from recognition by N-recognins, to destruction by AAA+proteases. *Biochim.Biophys.Acta.* **1823,** 83-91

Dougan, D.A., Mogk, A., and Bukau, B. (2002a) Protein folding and degradation in bacteria: to degrade or not to degrade? That is the question. *Cell Mol.Life Sci.* **59,** 1607-1616

Dougan, D.A., Reid, B.G., Horwich, A.L., and Bukau, B. (2002b) ClpS, a substrate modulator of the ClpAP machine. *Mol.Cell.* **9,** 673-683

Dougan, D.A., Truscott, K.N., and Zeth, K. (2010) The bacterial N-end rule pathway: expect the unexpected. *Mol.Microbiol.* **76,** 545-558

Ebhardt, H.A., Xu, Z., Fung, A.W., and Fahlman, R.P. (2009) Quantification of the post-translational addition of amino acids to proteins by MALDI-TOF mass spectrometry. *Anal.Chem.* **81,** 1937-1943

Erbse, A., Schmidt, R., Bornemann, T., Schneider-Mergener, J., Mogk, A., Zahn, R., Dougan, D.A., and Bukau, B. (2006) ClpS is an essential component of the N-end rule pathway in *Escherichia coli. Nature.* **439,** 753-756

Flynn, J.M., Neher, S.B., Kim, Y.I., Sauer, R.T., and Baker, T.A. (2003) Proteomic discovery of cellular substrates of the ClpXP protease reveals five classes of ClpX-recognition signals. *Mol.Cell.* **11,** 671-683

Fung, A.W., Ebhardt, H.A., Abeysundara, H., Moore, J., Xu, Z., and Fahlman, R.P. (2011) An alternative mechanism for the catalysis of peptide bond formation by L/F transferase: substrate binding and orientation. *J.Mol.Biol.* **409,** 617-629

Fung, A.W., Ebhardt, H.A., Krishnakumar, K.S., Moore, J., Xu, Z., Strazewski, P., and Fahlman, R.P. (2014a) Probing the Leucyl/Phenylalanyl tRNA Protein Transferase Active Site with tRNA Substrate Analogues. *Protein Pept.Lett.* **21,** 603-614

Fung, A.W., Leung, C.C., and Fahlman, R.P. (2014b) The determination of $tRNA^{Leu}$ recognition nucleotides for *Escherichia coli* L/F transferase. *RNA.* **20,** 1210-1222

Garg, R.P., Qian, X.L., Alemany, L.B., Moran, S., and Parry, R.J. (2008) Investigations of valanimycin biosynthesis: elucidation of the role of seryl-tRNA. *Proc.Natl.Acad.Sci.U.S.A.* **105,** 6543-6547

Gibbs, D.J., Bacardit, J., Bachmair, A., and Holdsworth, M.J. (2014) The eukaryotic N-end rule pathway: conserved mechanisms and diverse functions. *Trends Cell Biol.*

Gonda, D.K., Bachmair, A., Wunning, I., Tobias, J.W., Lane, W.S., and Varshavsky, A. (1989) Universality and structure of the N-end rule. *J.Biol.Chem.* **264,** 16700-16712

Graciet, E., Hu, R.G., Piatkov, K., Rhee, J.H., Schwarz, E.M., and Varshavsky, A. (2006) Aminoacyl-transferases and the N-end rule pathway of prokaryotic/eukaryotic specificity in a human pathogen. *Proc.Natl.Acad.Sci.U.S.A.* **103,** 3078-3083

Graciet, E., and Wellmer, F. (2010) The plant N-end rule pathway: structure and functions. *Trends Plant Sci.* **15,** 447-453

Guo, F., Esser, L., Singh, S.K., Maurizi, M.R., and Xia, D. (2002) Crystal structure of the heterodimeric complex of the adaptor, ClpS, with the N-domain of the AAA+ chaperone, ClpA. *J.Biol.Chem.* **277,** 46753-46762

Gur, E., Biran, D., and Ron, E.Z. (2011) Regulated proteolysis in Gram-negative bacteria--how and when?. *Nat.Rev.Microbiol.* **9,** 839-848

Hebecker, S., Arendt, W., Heinemann, I.U., Tiefenau, J.H., Nimtz, M., Rohde, M., Soll, D., and Moser, J. (2011) Alanyl-phosphatidylglycerol synthase: mechanism of substrate recognition during tRNA-dependent lipid modification in Pseudomonas aeruginosa. *Mol.Microbiol.* **80,** 935-950

Herman-Bachinsky, Y., Ryoo, H.D., Ciechanover, A., and Gonen, H. (2007) Regulation of the *Drosophila* ubiquitin ligase DIAP1 is mediated via several distinct ubiquitin system pathways. *Cell Death Differ.* **14,** 861-871

Hu, R.G., Sheng, J., Qi, X., Xu, Z., Takahashi, T.T., and Varshavsky, A. (2005) The N-end rule pathway as a nitric oxide sensor controlling the levels of multiple regulators. *Nature.* **437,** 981-986

Humbard, M.A., Surkov, S., De Donatis, G.M., Jenkins, L.M., and Maurizi, M.R. (2013) The N-degradome of *Escherichia coli*: limited proteolysis *in vivo* generates a large pool of proteins bearing N-degrons. *J.Biol.Chem.* **288,** 28913-28924

Hwang, C.S., Shemorry, A., and Varshavsky, A. (2010) N-terminal acetylation of cellular proteins creates specific degradation signals. *Science.* **327,** 973-977

Ichetovkin, I.E., Abramochkin, G., and Shrader, T.E. (1997) Substrate recognition by the leucyl/phenylalanyl-tRNA-protein transferase. Conservation within the enzyme family and localization to the trypsin-resistant domain. *J.Biol.Chem.* **272,** 33009-33014

Igarashi, K., and Kashiwagi, K. (2000) Polyamines: mysterious modulators of cellular functions. *Biochem.Biophys.Res.Commun.* **271,** 559-564

Iyer, L.M., Abhiman, S., Maxwell Burroughs, A., and Aravind, L. (2009) Amidoligases with ATP-grasp, glutamine synthetase-like and acetyltransferase-like domains: synthesis of novel metabolites and peptide modifications of proteins. *Mol.Biosyst.* **5,** 1636-1660

Jenal, U. (2009) The role of proteolysis in the Caulobacter crescentus cell cycle and development. *Res.Microbiol.* **160,** 687-695

Kaji, A., Kaji, H., and Novelli, G.D. (1965a) Soluble Amino Acid-Incorporating System. I. Preparation of the System and Nature of the Reaction. *J.Biol.Chem.* **240,** 1185-1191

Kaji, A., Kaji, H., and Novelli, G.D. (1965b) Soluble Amino Acid-Incorporating System. II. Soluble Nature of the System and the Characterization of the Radioactive Product. *J.Biol.Chem.* **240,** 1192-1197

Kawaguchi, J., Maejima, K., Kuroiwa, H., and Taki, M. (2013) Kinetic analysis of the leucyl/phenylalanyl-tRNA-protein transferase with acceptor peptides possessing different N-terminal penultimate residues. *FEBS Open Bio.* **3,** 252-255

Kim, H.K., Kim, R.R., Oh, J.H., Cho, H., Varshavsky, A., and Hwang, C.S. (2014) The N-terminal methionine of cellular proteins as a degradation signal. *Cell.* **156,** 158-169

Kwon, Y.T., Kashina, A.S., Davydov, I.V., Hu, R.G., An, J.Y., Seo, J.W., Du, F., and Varshavsky, A. (2002) An essential role of N-terminal arginylation in cardiovascular development. *Science.* **297,** 96-99

Kwon, Y.T., Kashina, A.S., and Varshavsky, A. (1999) Alternative splicing results in differential expression, activity, and localization of the two forms of arginyl-tRNA-protein transferase, a component of the N-end rule pathway. *Mol.Cell.Biol.* **19,** 182-193

Lee, M.J., Tasaki, T., Moroi, K., An, J.Y., Kimura, S., Davydov, I.V., and Kwon, Y.T. (2005) RGS4 and RGS5 are *in vivo* substrates of the N-end rule pathway. *Proc.Natl.Acad.Sci.U.S.A.* **102,** 15030-15035

Leibowitz, M.J., and Soffer, R.L. (1969) A soluble enzyme from *Escherichia coli* which catalyzes the transfer of leucine and phenylalanine from tRNA to acceptor proteins. *Biochem.Biophys.Res.Commun.* **36,** 47-53

Leibowitz, M.J., and Soffer, R.L. (1970) Enzymatic modification of proteins. III. Purification and properties of a leucyl, phenylalanyl transfer

ribonucleic acid protein transferase from *Escherichia coli. J.Biol.Chem.* **245,** 2066-2073

Leibowitz, M.J., and Soffer, R.L. (1971a) Enzymatic modification of proteins. VII. Substrate specificity of leucyl,phenylalanyl-transfer ribonucleic acid-protein transferase*. J.Biol.Chem.* **246,** 5207-5212

Leibowitz, M.J., and Soffer, R.L. (1971b) Modification of a specific ribosomal protein catalyzed by leucyl, phenylalanyl-tRNA: protein transferase*. Proc.Natl.Acad.Sci.U.S.A.* **68,** 1866-1869

Leu, N.A., Kurosaka, S., and Kashina, A. (2009) Conditional Tek promoter-driven deletion of arginyltransferase in the germ line causes defects in gametogenesis and early embryonic lethality in mice*. PLoS One.* **4,** e7734

Li, J., and Pickart, C.M. (1995) Binding of phenylarsenoxide to Arg-tRNA protein transferase is independent of vicinal thiols*. Biochemistry.* **34,** 15829-15837

Maglica, Z., Kolygo, K., and Weber-Ban, E. (2009) Optimal efficiency of ClpAP and ClpXP chaperone-proteases is achieved by architectural symmetry*. Structure.* **17,** 508-516

Moliere, N., and Turgay, K. (2009) Chaperone-protease systems in regulation and protein quality control in Bacillus subtilis*. Res.Microbiol.* **160,** 637-644

Ninnis, R.L., Spall, S.K., Talbo, G.H., Truscott, K.N., and Dougan, D.A. (2009) Modification of PATase by L/F-transferase generates a ClpS-dependent N-end rule substrate in *Escherichia coli. EMBO J.* **28,** 1732-1744

Potuschak, T., Stary, S., Schlogelhofer, P., Becker, F., Nejinskaia, V., and Bachmair, A. (1998) PRT1 of Arabidopsis thaliana encodes a component of the plant N-end rule pathway*. Proc.Natl.Acad.Sci.U.S.A.* **95,** 7904-7908

Rai, R., and Kashina, A. (2005) Identification of mammalian arginyltransferases that modify a specific subset of protein substrates*. Proc.Natl.Acad.Sci.U.S.A.* **102,** 10123-10128

Rai, R., Mushegian, A., Makarova, K., and Kashina, A. (2006) Molecular dissection of arginyltransferases guided by similarity to bacterial peptidoglycan synthases*. EMBO Rep.* **7,** 800-805

Rajagopala, S.V., Sikorski, P., Kumar, A., Mosca, R., Vlasblom, J., Arnold, R., Franca-Koh, J., Pakala, S.B., Phanse, S., Ceol, A., Hauser, R., Siszler, G., Wuchty, S., Emili, A., Babu, M., Aloy, P., Pieper, R., and Uetz, P. (2014) The binary protein-protein interaction landscape of *Escherichia coli. Nat.Biotechnol.* **32,** 285-290

Rao, H., Uhlmann, F., Nasmyth, K., and Varshavsky, A. (2001) Degradation of a cohesin subunit by the N-end rule pathway is essential for chromosome stability*. Nature.* **410,** 955-959

Rao, P.M., and Kaji, H. (1974) Utilization of isoaccepting leucyl-tRNA in the soluble incorporation system and protein synthesizing systems from *E.coli. FEBS Lett.* **43,** 199-202

Roy, H., and Ibba, M. (2008) RNA-dependent lipid remodeling by bacterial multiple peptide resistance factors*. Proc.Natl.Acad.Sci.U.S.A.* **105,** 4667-4672

Saha, S., and Kashina, A. (2011) Posttranslational arginylation as a global biological regulator*. Dev.Biol.* **358,** 1-8

Samsonova, N.N., Smirnov, S.V., Altman, I.B., and Ptitsyn, L.R. (2003) Molecular cloning and characterization of *Escherichia coli* K12 *ygjG* gene*. BMC Microbiol.* **3,** 2

Sauer, R.T., and Baker, T.A. (2011) AAA+ proteases: ATP-fueled machines of protein destruction*. Annu.Rev.Biochem.* **80,** 587-612

Scarpulla, R.C., Deutch, C.E., and Soffer, R.L. (1976) Transfer of methionyl residues by leucyl, phenylalanyl-tRNA-protein transferase*. Biochem.Biophys.Res.Commun.* **71,** 584-589

Scarpulla, R.C., and Soffer, R.L. (1978) Membrane-bound proline dehydrogenase from *Escherichia coli.* Solubilization, purification, and characterization*. J.Biol.Chem.* **253,** 5997-6001

Scarpulla, R.C., and Soffer, R.L. (1979) Regulation of proline dehydrogenase activity in Escherichia coli by leucyl-, phenylalanyl-tRNA:protein transferase*. J.Biol.Chem.* **254,** 1724-1725

Schmidt, R., Zahn, R., Bukau, B., and Mogk, A. (2009) ClpS is the recognition component for *Escherichia coli* substrates of the N-end rule degradation pathway*. Mol.Microbiol.* **72,** 506-517

Schuenemann, V.J., Kralik, S.M., Albrecht, R., Spall, S.K., Truscott, K.N., Dougan, D.A., and Zeth, K. (2009) Structural basis of N-end rule substrate

recognition in *Escherichia coli* by the ClpAP adaptor protein ClpS. *EMBO Rep.* **10,** 508-514

Shrader, T.E., Tobias, J.W., and Varshavsky, A. (1993) The N-end rule in *Escherichia coli*: cloning and analysis of the leucyl, phenylalanyl-tRNA-protein transferase gene *aat. J.Bacteriol.* **175,** 4364-4374

Siegele, D.A., and Kolter, R. (1993) Isolation and characterization of an *Escherichia coli* mutant defective in resuming growth after starvation. *Genes Dev.* **7,** 2629-2640

Soffer, R.L. (1973) Peptide acceptors in the leucine, phenylalanine transfer reaction. *J.Biol.Chem.* **248,** 8424-8428

Soffer, R.L. (1974) Aminoacyl-tRNA transferases. *Adv.Enzymol.Relat.Areas Mol.Biol.* **40,** 91-139

Soffer, R.L., Horinishi, H., and Leibowitz, M.J. (1969) The aminoacyl tRNA-protein transferases. *Cold Spring Harb.Symp.Quant.Biol.* **34,** 529-533

Soffer, R.L., and Savage, M. (1974) A mutant of *Escherichia coli* defective in leucyl, phenylalanyl-tRNA-protein transferase. *Proc.Natl.Acad.Sci.U.S.A.* **71,** 1004-1007

Sprangers, R., Gribun, A., Hwang, P.M., Houry, W.A., and Kay, L.E. (2005) Quantitative NMR spectroscopy of supramolecular complexes: dynamic side pores in ClpP are important for product release. *Proc.Natl.Acad.Sci.U.S.A.* **102,** 16678-16683

Stephani, K., Weichart, D., and Hengge, R. (2003) Dynamic control of Dps protein levels by ClpXP and ClpAP proteases in *Escherichia coli*. *Mol.Microbiol.* **49,** 1605-1614

Suto, K., Shimizu, Y., Watanabe, K., Ueda, T., Fukai, S., Nureki, O., and Tomita, K. (2006) Crystal structures of leucyl/phenylalanyl-tRNA-protein transferase and its complex with an aminoacyl-tRNA analog. *EMBO J.* **25,** 5942-5950

Tabor, C.W., and Tabor, H. (1985) Polyamines in microorganisms. *Microbiol.Rev.* **49,** 81-99

Tobias, J.W., Shrader, T.E., Rocap, G., and Varshavsky, A. (1991) The N-end rule in bacteria. *Science.* **254,** 1374-1377

Varshavsky, A. (2004) 'Spalog' and 'sequelog': neutral terms for spatial and sequence similarity. *Curr.Biol.* **14,** R181-3

Varshavsky, A. (2011) The N-end rule pathway and regulation by proteolysis. *Protein Sci.*

Vetting, M.W., S de Carvalho, L.P., Yu, M., Hegde, S.S., Magnet, S., Roderick, S.L., and Blanchard, J.S. (2005) Structure and functions of the GNAT superfamily of acetyltransferases. *Arch.Biochem.Biophys.* **433,** 212-226

Wagner, A.M., Fegley, M.W., Warner, J.B., Grindley, C.L., Marotta, N.P., and Petersson, E.J. (2011) N-terminal protein modification using simple aminoacyl transferase substrates. *J.Am.Chem.Soc.* **133,** 15139-15147

Wang, J., Hartling, J.A., and Flanagan, J.M. (1997) The structure of ClpP at 2.3 A resolution suggests a model for ATP-dependent proteolysis. *Cell.* **91,** 447-456

Wang, K.H., Oakes, E.S., Sauer, R.T., and Baker, T.A. (2008a) Tuning the strength of a bacterial N-end rule degradation signal. *J.Biol.Chem.* **283,** 24600-24607

Wang, K.H., Roman-Hernandez, G., Grant, R.A., Sauer, R.T., and Baker, T.A. (2008b) The molecular basis of N-end rule recognition. *Mol.Cell.* **32,** 406-414

Wang, K.H., Sauer, R.T., and Baker, T.A. (2007) ClpS modulates but is not essential for bacterial N-end rule degradation. *Genes Dev.* **21,** 403-408

Watanabe, K., Toh, Y., Suto, K., Shimizu, Y., Oka, N., Wada, T., and Tomita, K. (2007) Protein-based peptide-bond formation by aminoacyl-tRNA protein transferase. *Nature.* **449,** 867-871

Weichart, D., Querfurth, N., Dreger, M., and Hengge-Aronis, R. (2003) Global role for ClpP-containing proteases in stationary-phase adaptation of *Escherichia coli. J.Bacteriol.* **185,** 115-125

Weston, S.A., Camble, R., Colls, J., Rosenbrock, G., Taylor, I., Egerton, M., Tucker, A.D., Tunnicliffe, A., Mistry, A., Mancia, F., de la Fortelle, E., Irwin, J., Bricogne, G., and Pauptit, R.A. (1998) Crystal structure of the anti-fungal target N-myristoyl transferase. *Nat.Struct.Biol.* **5,** 213-221

Wickliffe, K.E., Leppla, S.H., and Moayeri, M. (2008) Killing of macrophages by anthrax lethal toxin: involvement of the N-end rule pathway. *Cell.Microbiol.* **10,** 1352-1362

Wolf, S.G., Frenkiel, D., Arad, T., Finkel, S.E., Kolter, R., and Minsky, A. (1999) DNA protection by stress-induced biocrystallization. *Nature.* **400,** 83-85

Wortham, B.W., Patel, C.N., and Oliveira, M.A. (2007) Polyamines in bacteria: pleiotropic effects yet specific mechanisms. *Adv.Exp.Med.Biol.* **603,** 106-115

Yamato, M., Umezawa, H., Sakata, N., Moriya, Y., and Hori, M. (1987) Valanimycin acts on DNA in bacterial cells. *J.Antibiot.(Tokyo).* **40,** 558-560

Zeth, K., Ravelli, R.B., Paal, K., Cusack, S., Bukau, B., and Dougan, D.A. (2002) Structural analysis of the adaptor protein ClpS in complex with the N-terminal domain of ClpA. *Nat.Struct.Biol.* **9,** 906-911

Zhang, W., Ntai, I., Kelleher, N.L., and Walsh, C.T. (2011) tRNA-dependent peptide bond formation by the transferase PacB in biosynthesis of the pacidamycin group of pentapeptidyl nucleoside antibiotics. *Proc.Natl.Acad.Sci.U.S.A.* **108,** 12249-12253

Zhou, Y., Larson, J.D., Bottoms, C.A., Arturo, E.C., Henzl, M.T., Jenkins, J.L., Nix, J.C., Becker, D.F., and Tanner, J.J. (2008) Structural basis of the transcriptional regulation of the proline utilization regulon by multifunctional PutA. *J.Mol.Biol.* **381,** 174-188

Chapter 2

Materials and Methods

A version of this chapter is published in:

Fung AW, Ebhardt HA, Abeysundara H, Moore J, Xu Z, and Fahlman RP (2011) An Alternative Mechanism for the Catalysis of Peptide Bond Formation by L/F Transferase: Substrate Binding and Orientation. *Journal of Molecular Biology.* 409 (4): 617-629.

Fung AW, Ebhardt HA, Krishnakumar KS, Moore J, Xu Z, Strazewski P, and Fahlman RP (2014) Probing the Leucyl/Phenylalanyl tRNA Protein Transferase Active Site with tRNA Substrate Analogues. *Protein and Peptide Letters*. 21 (7): 603-614.

Fung AW, Leung CC and Fahlman RP (2014) The Determination of $tRNA^{Leu}$ Recognition Nucleotides for *Escherichia coli* L/F transferase. *RNA*. 20 (8): 1210-1222.

2.1. Materials

Unless stated otherwise, all chemicals were purchased from Sigma-Aldrich. R-Cyano-4-hydroxycinnamic acid (CHCA), triethylamine (TEA), and bromoethane were purchased from Acros Organics.

2.2. Expression Vectors

2.2.1. L/F transferase

A clone of the wild type *E. coli* L/F transferase with an N-terminal 6× histidine tag in a pCA24N expression vector was obtained from the ASKA (-) strain collection (Kitagawa *et al.* 2005) from the National Institute of Genetics (Japan). Mutations to the wild type L/F transferase sequence were performed by site directed mutagenesis. For each point mutation the DNA oligo pairs (IDT, USA) listed in **Table 2-1** were used. All mutations were verified by DNA sequencing by the Applied Genomics Centre (Department of Medical Genetics, University of Alberta, Canada) or by the Eurofins MWG Operon (Huntsville, AL, USA).

2.2.2. Aminoacyl-tRNA Synthetases

A clone of an N-terminal 6× histidine tagged *E. coli* phenylalanyl-tRNA synthetase (PheRS) in a pET28a expression vector was a gift from Jack Szostak (Harvard Medical School). A clone of an N-terminal 6x histidine tagged *E. coli* leucyl-tRNA synthetase (LeuRS) in a pCA24N expression vector was obtained from the National Institute of Genetics

Table 2-1: List of primers used for L/F transferase mutagenesis.

Protein	5′-primer (forward)	3′-primer (reverse)
W49A	TGG TAT TTT TCC GGC GTT TTC TCC AGG CG	CGC CTG GAG AAA ACG CCG GAA AAA TAC CA
W111A	CGA AGA AGG AAC CGC GAT CAC GCG TGG CG	CGC CAC GCG TGA TCG CGG TTC CTT CTT CG
M144A	CTT GTC GGC GGT GCG TAC GGC GTG GCC	GGC CAC GCC GTA CGC ACC GCC GAC AAG
M144F	GAG CTT GTC GGC GGT TTC TAC GGC GTG GCC CAG	CTG GGC CAC GCC GTA GAA ACC GCC GAC AAG CTC
M144I	GAG CTT GTC GGC GGT ATC TAC GGC GTG GCC CAG	CTG GGC CAC GCC GTA GAT ACC GCC GAC AAG CTC
M144L	GAG CTT GTC GGC GGT TTG TAC GGC GTG GCC CAG	CTG GGC CAC GCC GTA CAA ACC GCC GAC AAG CTC
M144V	GAG CTT GTC GGC GGT GTG TAC GGC GTG GCC CAG	CTG GGC CAC GCC GTA CAC ACC GCC GAC AAG CTC
D186A	CGG TAA GCT TAT CGC CTG CCA GGT CCT TAA C	GTT AAG GAC CTG GCA GGC GAT AAG CTT ACC G
D186E	CGG TAA GCT TAT CGA ATG CCA GGT CCT TAA C	GTT AAG GAC CTG GCA TTC GAT AAG CTT ACC G
D186N	CGG TAA GCT TAT CAA CTG CCA GGT CCT TAA C	GTT AAG GAC CTG GCA GTT GAT AAG CTT ACC G
Q188A	GCT TAT CGA CTG CGC GGT CCT TAA CGA TC	GAT CGT TAA GGA CCG CGC AGT CGA TAA GC

All primer sequences are displayed from 5′ to 3′. DNA nucleotides for the mutated residue are underlined.

(Japan). Methionyl-tRNA synthetase (MetRS) was cloned into a pET28a plasmid vector between the NheI and NotI restriction sites which incorporates an N-terminal 6× histidine tag.

2.2.3. CCA adding enzyme

Cloned 6x histidine tagged nucleotidyl transferase (CCA adding enzyme) in a pET22b expression plasmid was a generous gift from Allen Weiner (University of Washington).

2.3. Protein Expression and Purification

The plasmids encoding the different proteins (L/F transferase and mutants, PheRS, LeuRS, MetRS, CCA adding enzyme, and T7 polymerase) were transformed into BL-21 DE3 cells. Transformed cells were grown to mid-log phase (~0.5 $O.D._{600nm}$) at 37 ºC in Luria Bertani media containing the appropriate selection antibiotic. Protein expression was induced by the addition of 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and the cultures were grown at 37 ºC for 4 hours or overnight. Cells were harvested by centrifugation and lysed by sonication in lysis buffer (20 mM NaH_2PO_4, 50 mM NaCl, 20 mM imidazole, 1 mM DTT, pH 7.4) containing 1 mM of the protease inhibitor phenylmethylsulphonyl fluoride (PMSF). Cleared lysates were passed through a 1 mL HisTrap FF Column (GE Healthcare). After extensive washing, the recombinant proteins were eluted in a linear gradient from lysis buffer solution to lysis buffer solution containing 1.0 M imidazole using an AKTA Prime FPLC (GE Healthcare). The eluted fractions were

analyzed by SDS-PAGE. The fractions containing the purified proteins were pooled together and dialyzed into storage buffer (50 mM Tris-Cl (pH 7.4), 100 mM NaCl, 1 mM DTT, 10% glycerol). Standard Bradford protein assay (Bio-Rad) and Bovine Serum Albumin (BSA) standards were used to determine the concentration of purified proteins. The purified proteins are diluted to 1 mg/mL, aliquoted into 100 μL per tube, flash frozen in liquid nitrogen, and stored at -80 ºC.

2.4. *In vitro* Transcription of tRNA

The DNA template for *in vitro* transcription of $tRNA^{Phe}$, $tRNA^{Leu}$, and $tRNA^{Met}$, and $tRNA^{Leu}$ hybrids were generated by primer extension using overlapping DNA oligonucleotides (IDT, USA) listed in **Table 2-2** and **2-3**, as previously described for $tRNA_{2A}^{Val}$ (Fahlman and Uhlenbeck 2004). Products of the primer extension were verified by agarose gel electrophoresis. Phenol/chloroform extraction and ethanol precipitation were then performed and the resulting pellet suspended in water for use in *in vitro* transcription reactions. *In vitro* transcription of tRNA was carried out as previously described (Sampson *et al.* 1987). Transcribed tRNAs were first concentrated by ethanol precipitation, resuspended in 5x RNA denaturing buffer (95% formamaide, 0.025% bromophenol blue, 5 mM EDTA), and separated by preparative 8% urea-polyacrylamide denaturing gel electrophoresis. The transcribed tRNA band was excised from the gel, extracted into 300 mM sodium acetate pH 5.0 by shaking overnight at 4 ºC, followed by 1-butanol extraction and ethanol precipitation. The

Table 2-2: List of primers used for wild type tRNA isoacceptors *in vitro* transcription.

tRNA	5′-primer (forward)	3′-primer (reverse)
$tRNA^{Phe}$ (GAA)	TAA TAC GAC TCA CTA TAG CCC GGA TAG CTC AGT CGG TAG AGC AGG **GGA TTG AAA ATC C**	TGG TGC CCG GAC TCG GAA TCG AAC CAA GGA CAC GG**G GAT TTT CAA TCC**
$tRNA^{Leu}$ (CAA)	TAA TAC GAC TCA CTA TAG CCG AAG TGG CGA AAT CGG TAG ACG CA**G TTG ATT CAA AAT C**	TGG TGC CGA AGG CCG GAC TCG AAC CGG CAC GTA TTT CTA CGG **TTG ATT TTG AAT CAA C**
$tRNA^{Leu}$ (CAG)	TAA TAC GAC TCA CTA TAG CGA AGG TGG CGG AAT TGG TAG ACG CGC TAG **CTT CAG GTG TT**	TGG TGC GAG GGG GGG GAC TTG AAC CCC CAC GTC CGT AAG AAC ACT **AAC ACC TGA AG**
$tRNA^{Leu}$ (GAG)	TAA TAC GAC TCA CTA TAG CCG AGG TGG TGG AAT TGG TAG ACA CGC **TAC CTT GAG GTG G**	TGG TAC CGA GGA CGG GAC TTG AAC CCG TAA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**
$tRNA^{Leu}$ (UAA)	TAA TAC GAC TCA CTA TAG CCC GGA TGG TGG AAT CGG TAG ACA CAA **GGG ATT TAA AAT C**	TGG TAC CCG GAG CGG GAC TTG AAC CCG CAC AGC GCG AAC GCC GAG G**GA TTT TAA ATC CC**
$tRNA^{Leu}$ (UAG)	TAA TAC GAC TCA CTA TAG CGG GAG TGG CGA AAT TGG TAG ACG CAC **CAG ATT TAG GTT C**	TGG TGC GGG AGG CGA GAC TTG AAC TCG CAC ACC TTG CGG CGC CA**G AAC CTA AAT CTG**
$tRNA^{Met}$ (CAU) *ilex*	TAA TAC GAC TCA CTA TAG GCC CCT TAG CTC AGT GGT TAG AGC AGG **CGA CTC ATA ATC G**	TGG TGG CCC CTG CTG GAC TTG AAC CAG CGA CCA AG**C GAT TAT GAG TCG**
$tRNA^{Met}$ (CAU) *metT*	TAA TAC GAC TCA CTA TAG GCT ACG TAG CTC AGT TG**G TTA GAG CAC ATC ACT CAT A**	TGG TGG CTA CGA CGG GAT TCG AAC CTG TGA CCC CAT CAT **TAT GAG TGA TGT GCT CTA AC**

All primer sequences are displayed from 5′ to 3′. T7 promoter sites are underlined and overlapped regions are bolded.

Table 2-3: List of primers used for tRNA^{Leu} hybrid *in vitro* transcription.

Construct	tRNA	5′-primer (forward)	3′-primer (reverse)
1	tRNA^{Leu} (GAG) U72C	TAA TAC GAC TCA CTA TAG CCG AGG TGG TGG AAT TGG TAG ACA CGC **TAC CTT GAG GTG G**	TGG T**G**C CGA GGA CGG GAC TTG AAC CCG TAA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**
2	tRNA^{Leu} (GAG) U68C	TAA TAC GAC TCA CTA TAG CCG AGG TGG TGG AAT TGG TAG ACA CGC **TAC CTT GAG GTG G**	TGG TAC CG**G** GGA CGG GAC TTG AAC CCG TAA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**
3	tRNA^{Leu} (GAG) G4A, C69U	TAA TAC GAC TCA CTA TAG CC**A** AGG TGG TGG AAT TGG TAG ACA CGC **TAC CTT GAG GTG G**	TGG TAC C**A**A GGA CGG GAC TTG AAC CCG TAA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**
4	tRNA^{Leu} (GAG) G4A, U68C, C69U	TAA TAC GAC TCA CTA TAG CC**A** AGG TGG TGG AAT TGG TAG ACA CGC **TAC CTT GAG GTG G**	TGG TAC C**AG** GGA CGG GAC TTG AAC CCG TAA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**
5	tRNA^{Leu} (GAG) G4A, U68C, C69U, U72C	TAA TAC GAC TCA CTA TAG CC**A** AGG TGG TGG AAT TGG TAG ACA CGC **TAC CTT GAG GTG G**	TGG T**G**C C**AG** GGA CGG GAC TTG AAC CCG TAA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**
6	tRNA^{Leu} (GAG) C3G, G70C	TAA TAC GAC TCA CTA TAG C**G**G AGG TGG TGG AAT TGG TAG ACA CGC **TAC CTT GAG GTG G**	TGG TAC **G**GA GGA CGG GAC TTG AAC CCG TAA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**
7	tRNA^{Leu} (GAG) C3G, G70C, U72C	TAA TAC GAC TCA CTA TAG C**G**G AGG TGG TGG AAT TGG TAG ACA CGC **TAC CTT GAG GTG G**	TGG T**G**C **G**GA GGA CGG GAC TTG AAC CCG TAA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**
8	tRNA^{Leu} (GAG) C3G, U68C, G70C	TAA TAC GAC TCA CTA TAG C**G**G AGG TGG TGG AAT TGG TAG ACA CGC **TAC CTT GAG GTG G**	TGG TAC **G**G**G** GGA CGG GAC TTG AAC CCG TAA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**

9	tRNA^{Leu} (GAG) C3G, G4A, C69U, G70C	TAA TAC GAC TCA CTA TAG C**GA** AGG TGG TGG AAT TGG TAG ACA CGC **TAC CTT GAG GTG G**	TGG TAC **GA**A GGA CGG GAC TTG AAC CCG TAA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**
10	tRNA^{Leu} (GAG) C3G, G4A, U68C, C69U, G70C	TAA TAC GAC TCA CTA TAG C**GA** AGG TGG TGG AAT TGG TAG ACA CGC **TAC CTT GAG GTG G**	TGG TAC **GAG** GGA CGG GAC TTG AAC CCG TAA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**
11	tRNA^{Leu} (GAG) C3G, G4A, U68C, C69U, G70C, U72C	TAA TAC GAC TCA CTA TAG C**GA** AGG TGG TGG AAT TGG TAG ACA CGC **TAC CTT GAG GTG G**	TGG T**G**C **GAG** GGA CGG GAC TTG AAC CCG TAA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**
12	tRNA^{Leu} (GAG) U11C, A24G	TAA TAC GAC TCA CTA TAG CCA AGG TGG **C**GG AAT TGG TAG AC**G** CGC **TAC CTT GAG GTG G**	TGG TAC CGA GGA CGG GAC TTG AAC CCG TAA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**
13	tRNA^{Leu} (GAG) A49G, U65C	TAA TAC GAC TCA CTA TAG CCG AGG TGG TGG AAT TGG TAG ACA CGC **TAC CTT GAG GTG G**	TGG TAC CGA GG**G** CGG GAC TTG AAC CCG **C**AA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**
14	tRNA^{Leu} (GAG) C3G, G4A, U11C, A24G, U68C, C69U, G70C, U72C	TAA TAC GAC TCA CTA TAG C**GA** AGG TGG **C**GG AAT TGG TAG AC**G** CGC **TAC CTT GAG GTG G**	TGG T**G**C **GAG** GGA CGG GAC TTG AAC CCG TAA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**
15	tRNA^{Leu} (GAG) C3G, G4A, U11C, A24G, A49G, U65C, U68C, C69U, G70C, U72C	TAA TAC GAC TCA CTA TAG C**GA** AGG TGG **C**GG AAT TGG TAG AC**G** CGC **TAC CTT GAG GTG G**	TGG T**G**C **GAG** GG**G** CGG GAC TTG AAC CCG **C**AA GCC CTA TTG GGC ACT A**CC ACC TCA AGG TA**

16	$tRNA^{Leu}$ (CAG) G3C, C70G	TAA TAC GAC TCA CTA TAG C**C**A AGG TGG CGG AAT TGG TAG ACG CGC TAG **CTT CAG GTG TT**	TGG TGC **C**AG GGG GGG GAC TTG AAC CCC CAC GTC CGT AAG AAC ACT **AAC ACC TGA AG**
17	$tRNA^{Leu}$ (CAG) A4G, C68U, U69C, C72U	TAA TAC GAC TCA CTA TAG CG**G** AGG TGG CGG AAT TGG TAG ACG CGC TAG **CTT CAG GTG TT**	TGG T**A**C G**GA** GGG GGG GAC TTG AAC CCC CAC GTC CGT AAG AAC ACT **AAC ACC TGA AG**
18	$tRNA^{Leu}$ (CAG) G3C, A4G, C68U, U69C, C70G, C72U	TAA TAC GAC TCA CTA TAG C**CG** AGG TGG CGG AAT TGG TAG ACG CGC TAG **CTT CAG GTG TT**	TGG T**A**C **CGA** GGG GGG GAC TTG AAC CCC CAC GTC CGT AAG AAC ACT **AAC ACC TGA AG**

All primer sequences are displayed from 5′ to 3′.

T7 promoter sites are underlined and overlapped regions are bolded.

Mutated residues are both underlined and bolded.

concentration of purified, unmodified tRNAs were determined by measuring absorbance at 260 nm.

2.5. Stable Isotope Labeling of Peptides

The substrate and product peptides (REPGLCTWQSLR, FREPGLCTWQSLR, LREPGLCTWQSLR, and MREPGLCTWQSLR) were purchased from the Institute for Biomolecular Design (University of Alberta, Canada). Peptide stock solutions were made and their absolute concentration was determined by amino acid analysis (Institute for Biomolecular Design). To generate a peptide pair of identical chemical composition that differ in mass by five Daltons, the substrate and product peptides were alkylated with either bromoethane or deuterated (d5)-bromoethane as has been previously described (Hale *et al.* 1996, Hale *et al.* 2004). To a 30 μL solution containing 250 μM peptide, 1 μL 100 mM tris (2-carboxyethyl) phosphine (TCEP) was added. The sample was incubated at 37 ºC for 10 minutes. Then, 100 μL acetonitrile, 1 μL triethylamine and 7 μL bromoethane or (d_5)-bromoethane were added and the sample was incubated at 37 ºC for an additional 30 minutes. The sample was repeatedly dried by speed vacuum and suspended in water. The final pellet was suspended in water. The quantitative conversions of the final heavy and light peptides were verified by MALDI-ToF mass spectrometry.

2.6. Puromycin and rA-Phe-amide Inhibitors Preparation

Puromycin (3'-deoxy-*N,N*-dimethyl-3'-[(*O*-methyl-L-tyrosyl) amino] adenosine, from Sigma-Aldrich) and rA-Phe-amide (3'-deoxy-3'-[(L-phenylalanyl) amino] adenosine, synthesized by Dr. Kollappillil Krishnakumar and Dr. Peter Strazewski from the Université de Lyon as part of research collaboration) were dissolved in distilled water. The stock concentrations of puromycin and rA-Phe-amide were determined by first measuring the UV absorbance of each at 260 nm and then calculated using the molar extinction coefficient of puromycin (15 400 $M^{-1}cm^{-1}$) and rA-Phe-amide (13 200 $M^{-1}cm^{-1}$).

2.7. L/F transferase Activity Assay

2.7.1. With Varying Peptide Substrate Concentrations

The peptide bond formation reaction was modified from the original procedure described by Ebhardt *et al.* (Ebhardt *et al.* 2009). A major alteration is the use of an internal standard peptide to quantify product formation as opposed to substrate disappearance. Additionally, the reaction set-up was optimized for use with multi-channel pipettors enabling eight reactions to be performed in parallel. All reactions were performed at 37 ºC in 1× reaction buffer (50 mM Hepes pH 7.5, 50 mM KCl, 15 mM $MgCl_2$). In each strip of 8 tubes (Corning Thermowell Gold PCR) was a 90 μL tRNA pre-charging reaction containing 1× reaction buffer (2 mM ATP, 0.2 mM CTP, 2 mM β-mercaptoethanol, 2 mM phenylalanine, 2 μM $tRNA^{Phe}$, 1 μL of 1 mg/mL purified CCA adding

enzyme, 1 μL of 1 mg/mL purified PheRS). These tubes also contained light-labeled substrate peptide and heavy-labeled product standard peptides in equimolar concentrations ranging from 0.30 to 14.78 μM. These reaction mixtures were incubated for 7 minutes at 37 ºC to facilitate aminoacylation of the tRNA. The reactions were initiated by the addition of wild-type or mutant L/F transferase (final concentration of 3.8 μM in **Chapter 3** and **5** and 7.6 μM in **Chapter 4** unless stated otherwise) to the samples containing the aminoacylated tRNAs and peptide substrate. At increasing time points (2 to 90 minutes), 5 μL of the reactions were added to 5 μL of quench solution (0.01 μg/μL BSA, 10% acetonitrile, 2% trifluoracetic acid). For analysis, 10 μL matrix solution (saturated R-cyano-4-hydroxycinnamic acid in 50% acetonitrile and 0.2% TFA) was added to the quenched reaction aliquots and 1 μL of the mixture was spotted in duplicate on a MALDI-ToF sample plate. The spectra were collected on a Bruker Daltonics Ultraflex MALDI-ToF/ToF at the Institute for Biomolecular Design (University of Alberta, Canada).

2.7.2. With Puromycin or rA-Phe-amide Inhibition

To measure puromycin inhibition on L/F transferase, the final concentration of reagents and enzymes were identical to the conditions described above with the exception of L/F transferase being pre-incubated at 37 °C for 7 minutes with varying concentrations of inhibitors (0, 0.24, 0.48, and 0.95 mM unless stated otherwise). The L/F transferase (wild-type or mutants)–puromycin mixture was then added to the reaction

mixture containing aa-tRNAs and substrate peptides, and the reaction mixtures were sampled and analyzed as described in **Section 2.7.1.** Analysis of rA-Phe-amide was performed similarly to puromycin treatments.

2.7.3. With Varying tRNA Substrate Concentrations

To measure L/F transferase activity with varying tRNA concentrations, *in vitro* transcribed tRNA ($tRNA^{Phe}$, $tRNA^{Leu}$, $tRNA^{Met}$, and $tRNA^{Leu}$ hybrids) were first re-folded at 65 ºC in 5 mM NaOAc and then slow cooled to room temperature. Then the final concentration of reagents and enzymes were identical to the conditions described above except that the light-labeled substrate peptide (REPGLCTWQSLR) and heavy-labeled standard product peptide (LREPGLCTWQSLR, FREPGLCTWQSLR, or MREPGLCTWQSLR) were fixed at a final concentration of 5.91 μM, meanwhile the re-folded tRNA final concentrations range from 1.25 to 50 μM. The reaction mixtures were incubated for 7 minutes at 37 ºC to facilitate the aminoacylation of the tRNAs. The reactions were initiated by the addition of wild-type L/F transferase to a final concentration of 3.8 μM. The reaction mixtures were sampled and analyzed as described above.

2.7.4. Competition with Uncharged tRNA

To measure whether the presence of uncharged tRNA (product of the reaction) inhibits L/F transferase, competition assays were performed. In a leucylation assay, first 5 μM of $tRNA^{Leu}$ (CAG) was aminoacylated by

LeuRS in a reaction mixture as described in **Section 2.7.3** for 7 minutes at 37 ºC. Meanwhile at the same time, 3.8 μM of L/F transferase was pre-incubated with 0, 25, 50, or 75 μM of uncharged $tRNA^{Phe}$ (GAA) for 7 minutes at 37 ºC. The L/F transferase-uncharged tRNA mixture is then mixed with the leucylated mixture, and time points were sampled and analyzed as described in **Section 2.7.1.** Alternatively in a phenylalanylation assay with $tRNA^{Phe}$ (GAA), uncharged $tRNA^{Leu}$ (CAG) was used instead.

2.8. Kinetic Data and Curve-fit Analysis

To quantify product formation, the ratio of area under the peak of the product peptide (ex. light product peptide *m/z* 1620) to area under the peak of the internal standard peptide (ex. heavy stand product peptide *m/z* 1625) was used to calculate the concentration of product formed at a specific time point at a given substrate concentration. A graph of product concentration against time was plotted for each experiment (varying substrate concentrations). A linear tangent line was drawn manually between time zero and the linear portion of the product concentration against time curve, where at least three time points were used in addition to time zero. The slope of the linear tangent line is the value for the initial rate of product formation for that specific substrate concentration. The initial rates of product formation calculated were then plotted against peptide substrate (**Chapter 3** and **4**) or tRNA substrate (**Chapter 5**) concentrations. GraphPad Prism Version 5.02 (GraphPad Software) was

used to calculate the apparent kinetic parameters (apparent K_M and apparent k_{cat}) by fitting the data to a non-linear regression analysis and defining the enzyme concentration. This method requires at least three data points at the saturating part of the Michaelis-Menten curve. Although the final values reported were derived from GraphPad Prism, primary or Lineweaver-Burk plots (reciprocal of initial rates of product formation against reciprocal of substrate concentration) were also generated to confirm the apparent kinetic parameters calculations.

The data in **Chapter 4** were further analyzed to determine the inhibition models using non-linear regression analysis by GraphPad Prism Version 5.02 (GraphPad Software). Of the following models tested (competitive, non-competitive, uncompetitive, and mixed), the data best fits to a non-competitive inhibition model. Subsequent curve-fit analyses were performed using the non-competitive inhibition option in the GraphPad Prism software. The apparent inhibition constants (K_i) were calculated by the GraphPad Prism software by defining the inhibitor concentrations. Primary or Lineweaver-Burk plots was generated to confirm the mode of inhibition. It is observed that wild-type L/F transferase best fits to a mixed competitive/ non-competitive inhibition model by the secondary plot method. Although the final kinetic parameter and inhibition constant values reported derived from GraphPad Prism, secondary plots (slope of Lineweaver-Burk plot against inhibitor

concentration) were also generated to confirm the apparent K_i calculations (x-intercept).

2.9. Radiolabeling tRNA and Aminoacylation Assay

To ensure that the differences in L/F transferase product formation rates are due to RNA structure and not due to reduced aminoacylation, percent aminoacylation were tested for each tRNA constructs. Aminoacylation assay has been modified as previously described (Wolfson and Uhlenbeck 2002). Briefly, 2 μM purified *in vitro* transcribed tRNAs were folded by heating for 3 min at 65 ºC in 10 mM $MgCl_2$. To ^{32}P-labeled the tRNAs at the 3' terminal internucleotide linkage, the folded tRNAs were added to a 100 μL reaction containing 50 mM glycine-HCl (pH 9.0), 50 μM $NaPP_i$, 2 μM [α-^{32}P] ATP [3,000 Ci/mmol], 0.06 μg/μL CCA adding enzyme, and 80 units of RNase OUT and incubate at 37 °C for 5 minutes. 2 μL of 10 units/mL Yeast PPase and 2 μM CTP were added, and the reaction mixture was incubated for an additional 2 minutes before quenching by phenol/chloroform extraction. After ethanol precipitation, 3'-^{32}P-labeled tRNAs were purified by pre-equilibrated desalting column (Thermo Fisher 7k MWCO). Aminoacylation levels were determined in a 25 μL reaction containing 1x Aminoacylation Buffer (50mM HEPES pH 7.5, 30 mM KCl, 15 mM $MgCl_2$, 2 mM ATP, 0.5 mM DTT), 1 mM amino acid (leucine, phenylalanine, or methionine), 0.8 μM 3'-^{32}P-labeled tRNA (isoaccepting or hybrid species), and 1 μL of 1 mg/mL aa-tRNA synthetase (LeuRS, PheRS, or MetRS) at 37 ºC. Aliquots were taken at

specific time points (0, 0.5, 1, 2, 3, 4, 5, 7, 10, and 20 minutes) and quenched on ice with 200 mM sodium acetate pH 5.0 containing 1 unit/μL of nuclease S1 or P1. The quenched aliquots were kept on ice until the aminoacylation time course was completed and then incubated for 10 minutes at room temperature to digest the tRNA into AMP and aminoacyl-AMP (aa-AMP). 1 μL of the digestion reaction was spotted on a 9-cm prewashed polyethylenimine-cellulose plates, and the AMP and AMP-AA were separated by TLC in glacial acetic acid/1M NH_4Cl/H_2O (5:10:85). The radioactivity was measured by PhosphorImager (Molecular Dynamics) and quantified by ImageQuant (Version 5.2). Percent aminoacylation is calculated from the fraction of aa-AMP over total (aa-AMP + AMP).

2.10. References

Ebhardt, H.A., Xu, Z., Fung, A.W., and Fahlman, R.P. (2009) Quantification of the post-translational addition of amino acids to proteins by MALDI-TOF mass spectrometry. *Anal.Chem.* **81,** 1937-1943

Fahlman, R.P., and Uhlenbeck, O.C. (2004) Contribution of the esterified amino acid to the binding of aminoacylated tRNAs to the ribosomal P- and A-sites. *Biochemistry.* **43,** 7575-7583

Hale, J.E., Butler, J.P., and Pourmand, R.R. (1996) Analysis of cysteine residues in peptides and proteins alkylated with volatile reagents. *Amino Acids.* **10,** 243-252

Hale, J.E., Butler, J.P., Gelfanova, V., You, J.S., and Knierman, M.D. (2004) A simplified procedure for the reduction and alkylation of cysteine residues in proteins prior to proteolytic digestion and mass spectral analysis. *Anal.Biochem.* **333,** 174-181

Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., and Mori, H. (2005) Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research. *DNA Res.* **12,** 291-299

Sampson, J.R., Sullivan, F.X., Behlen, L.S., DiRenzo, A.B., and Uhlenbeck, O.C. (1987) Characterization of two RNA-catalyzed RNA cleavage reactions. *Cold Spring Harb.Symp.Quant.Biol.* **52,** 267-275

Wolfson, A.D., and Uhlenbeck, O.C. (2002) Modulation of $tRNA^{Ala}$ identity by inorganic pyrophosphatase. *Proc.Natl.Acad.Sci.U.S.A.* **99,** 5965-5970

Chapter 3

The Functional Role of D186 –

Substrate Binding and Orientation

A version of this chapter is published in:

Fung AW, Ebhardt HA, Abeysundara H, Moore J, Xu Z, and Fahlman RP (2011) An Alternative Mechanism for the Catalysis of Peptide Bond Formation by L/F Transferase: Substrate Binding and Orientation. *Journal of Molecular Biology.* 409 (4): 617-629.

3.1. Introduction

The non-ribosomal addition of amino acids to proteins was first described almost 50 years ago in *E. coli* (Kaji *et al.* 1965a, Kaji *et al.* 1965b). This process is also described as tRNA-dependent non-ribosomal peptide bond formation. It is an acyl transfer reaction where the aminoacyl moiety transfers from the 3' end of an aa-tRNA substrate to an acceptor substrate resulting in an amide linkage (or peptide bond).

Recent availability of X-ray crystal structures of L/F transferase has provided new opportunities for investigating the molecular mechanisms of this enzyme (Suto *et al.* 2006, Dong *et al.* 2007, Watanabe *et al.* 2007). Co-crystal structures with bound substrates and products have resulted in a recently proposed mechanism for dynamics and catalysis (Watanabe *et al.* 2007). The superposition of the substrate analogue phenylalanyl adenosine (rA-Phe) and the product peptide (FRYLG) complex shows significant changes in the positioning for the following three key residues: Q188, W49, and W111 (**Figure 3-1**). Q188 was proposed to be a catalytic residue as described below. W49 and W111 have been suggested to have a role in recognizing the aa-tRNA substrate via π-π base stacking interactions and alanine substitution mutations to these two amino acids result in nearly complete inactivity (Suto *et al.* 2006, Watanabe *et al.* 2007). However since there is no crystal structure solved for L/F transferase in complex with tRNA bound or an intermediate analogue bound and given that there are discrepancies in the structures of tRNA

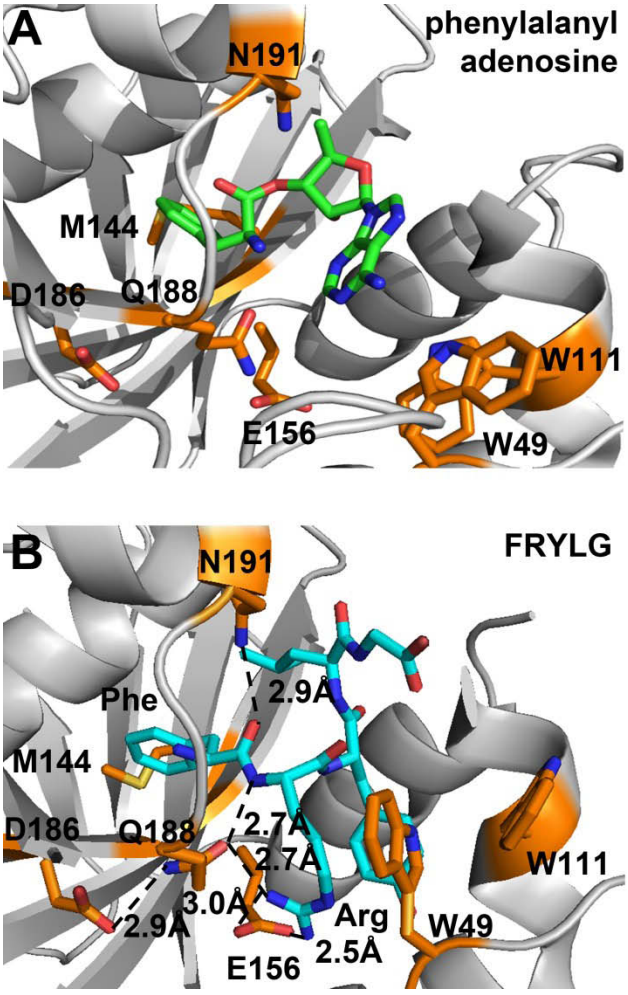

Figure 3-1: X-ray crystal structures of L/F transferase active site. A) An X-ray crystal structure of L/F transferase with a phenylalanyl adenosine bound (rA-Phe, green, sticks). rA-Phe is an aa-tRNA substrate analogue. The amino acids under investigation are highlighted in orange. Note that D186 is not in contact with the bound substrate analogue and may function in positioning key residues during catalysis. This image was prepared from a protein data bank file, PDB ID: 2Z3K, using PyMOL Molecular Graphics System, Version 1.3, Schrödinger, LLC. **B)** An X-ray crystal structure of L/F transferase with a bound product peptide (FRYLG, cyan, sticks). This structure reveals extensive hydrogen bonds that stabilize the formed product (key residues are highlighted in orange). D186 is again not in contact to the bound product but is hydrogen bonded to Q188 (PDB ID: 2Z3N).

analogues puromycin and rA-Phe (see **Chapter 4**), the molecular insights derived from these structures remain somewhat uncertain.

An initial protein-based acid/base catalytic mechanism has been proposed for L/F transferase where the enzyme is directly involved in the bond breaking and bond forming chemistries. Q188, activated by D186 via an electron-relay system, acts as a general base and attracts a proton from the α-amino group of the N-terminal Arno luncg substrate facilitating the nucleophilic attack to the carbonyl carbon of the aa-tRNA substrate (Watanabe *et al.* 2007). Through hydrogen bonding, N191 has been proposed to not only enhance the electrophilicity of the carbonyl carbon of aa-tRNA, but also stabilizes the tetrahedral oxyanion intermediate formed (Watanabe *et al.* 2007). The aminoacyl transfer completes with the protonation of the 3'-O of deacylated tRNA from Q188, and the products (product peptide and deacylated tRNA) are released. This proposed mechanism is similar to the classic reverse acylation step of proteolysis observed in serine proteases (Watanabe *et al.* 2007).

This proposal of glutamine acting as a general base is unconventional. Previous investigation reported that the Q188A mutant retains some enzymatic activity which was rationalized with a model where a water molecule replaces Q188 and it is similarly activated by D186 (Watanabe *et al.* 2007). To obtain further insight into the mechanism of L/F transferase, we focused our investigations on the role of D186, the activating amino acid involved in the proposed protein-based

acid/base mechanism (Watanabe *et al.* 2007). It has been suggested that D186 is essential for catalysis as it is absolutely conserved in L/F transferase sequences from different prokaryotic species (Ichetovkin *et al.* 1997, Suto *et al.* 2006, Watanabe *et al.* 2007) and the D186A mutation was reported to be completely inactive (Watanabe *et al.* 2007). Examination of the X-ray crystal structures of L/F transferase with a bound aa-tRNA substrate analogue rA-Phe (**Figure 3-1A**) or a bound product peptide (**Figure 3-1B**) reveals that D186 does not make contact with either substrate or product. What is apparent upon examination of the different X-ray crystal structures is that D186 may be involved in the positioning of Q188, which is in direct contact with the product peptide and, by extrapolation, is likely to be involved in binding the substrate peptide. If D186 is crucial for accurate positioning of Q188, then inactivation of the enzyme by the D186A mutation may deform the active site and impair enzymatic activity as a result of suboptimal substrate positioning.

Here we determine whether the size, hydrogen bonding property or carboxylic acid functionality of D186 is critical for L/F transferase activity. We have generated two additional mutants of L/F transferase (D186N and D186E) to compare with the wild-type enzyme and the previously reported D186A mutant. We purified and assayed the enzymatic activity of these enzymes to further characterize the role of this highly conserved residue in L/F transferase-catalyzed tRNA-dependent peptide bond formation.

Kinetic analysis was performed using a modified matrix assisted laser desorption/ionization time-of-flight (MALDI-ToF) mass spectrometry (MS) procedure that we have previously described (Ebhardt *et al.* 2009). To complement the analysis, we have also analyzed the previously reported mutations of Q188A, W49A, W111A, and M144A by quantitative MALDI-ToF MS for appropriate comparisons using the same analytical procedure. Our results demonstrate that despite significant loss of enzymatic activity for all the mutations investigated, some product formation is always observed. With our presented kinetic data and existing X-ray crystal structures, we propose an alternative catalytic mechanism. The roles of D186 and Q188, similarly to other amino acids examined, are for proper substrate binding and orientation and are not directly involved in the chemistry of the reaction as previously proposed. This alternative mechanism mirrors the proton-shuttling mechanism that has been described for the ribosome (Weinger and Strobel 2006, Beringer and Rodnina 2007, Fung *et al.* 2011).

3.2. Results

3.2.1. L/F Transferase D186 Mutations.

The absolute conservation of D186 in L/F transferase sequences from various prokaryotic species and the observation that the D186A mutant resulted in a completely inactive enzyme have led to the suggestion that it is essential for catalysis (Suto *et al.* 2006, Watanabe *et al.* 2007). Inactivation of L/F transferase by the original D186A mutation

can be rationalized in more than one way. On one hand D186 may have a role in positioning Q188 into a specific stereochemical arrangement. If so the reduced volume of the side chain of the D186A mutant could render the enzyme inactive by repositioning Q188 and other key amino acids. D186 may, on the other hand, have a role utilizing its negatively charged carboxyl functional group. To distinguish between the different possibilities for the inactivation of D186A mutation and to investigate the critical role of D186, we generated additional D186 point mutants. Both D186E and D186N L/F transferase mutant proteins were expressed, purified, and enzymatically assayed to be compared with both wild-type and the D186A mutant.

3.2.2. L/F Transferase Activity Assay.

The application of stable isotope labelling for the use of internal standards for mass spectrometry analysis has long been known (Desiderio and Kai 1983). To quantify the activity of the wild-type and mutant L/F transferase enzymes, we applied a modified MALDI-ToF MS based assay as previously described (Ebhardt *et al.* 2009). This procedure has been demonstrated to be much faster to perform and more sensitive than conventional autoradiography methods (Kuno *et al.* 2003) and more specific than trichloroacetic acid (TCA) precipitation protocols (Abramochkin and Shrader 1995). Also using mass spectrometry for product detection validates the existence of product even when trace amounts are formed. The modified procedure used in this manuscript

quantifies product peptide appearance by the application of stable isotope labelling and internal peptide standards. Chemically synthesized substrate (REPGLC*TWQSLR) and product (FREPGLC*TWQSLR) peptides were alkylated at the internal cysteine with either bromoethane or deuterated [$^{2}H_{5}$]-bromoethane to generate 'light' and 'heavy' pairs of substrate and product peptides. The incorporation of a light substrate peptide into the L/F transferase reaction results in the enzymatic formation of a light product peptide. Prior to analysis the heavy product peptide is added to the sample as an internal standard for quantification. As the light and heavy peptide pair is identical in their physical and chemical properties, they co-crystallize with matrix and ionize identically during MALDI-ToF analysis. The amount of light product peptide formed during the enzymatic reaction is quantified by measuring the relative ratio of the ion intensity of the light product peptide to the ion intensity of the heavy standard product peptide (known amount) that was added for mass spectrometry analysis. Our experimental data reveals no difference in the data when a standard product peptide is included in the reaction mixture or added with the reaction quench (data not shown). This suggests that there is negligible product peptide inhibition of the L/F transferase enzyme under our (this study) and other (Wagner *et al.* 2011) experimental conditions.

Raw data from mass spectrometry analysis of samples from three different time points of the wild-type L/F transferase reaction are shown in

Figure 3-2A. In the foreground is the spectrum from a two minute time point. From the data it can be seen that the light substrate is observed as a single peak at a mass to charge (*m/z*) of 1473, while the light product and heavy standard product are observed as a doublet at an *m/z* of 1620 and 1625 respectively. Even at two minutes a measurable amount of product (*m/z* 1620) has formed. As the reaction proceeds, it is readily apparent from the mass spectra at 20 and 60 minutes that there is a decrease in the relative intensity of light substrate peptide (*m/z* 1473) and a concurrent increase in intensity of the light product peptide at an *m/z* of 1620. The change in mass to charge (Δ *m/z*) of 147 between the light substrate peptide and the light product peptide reflects the mass of phenylalanine addition.

The ratio of intensities between light product (*m/z* 1620) and heavy standard product (*m/z* 1625) of known concentration (1.5 μM in the depicted sample data) enables the absolute quantification of product present in the sample. The entire procedure is not limited to the use of light peptide as substrate and heavy peptide as standard product. Their use can be reversed, where a heavy substrate peptide and a light standard product is used for quantification. The data in **Figure 3-2B** shows that there is no difference in the initial rate of product formation when either light substrate or heavy substrate is used in a reaction and this alternation of light and heavy peptides as substrates is performed during replicate analysis. The final reported values are obtained from

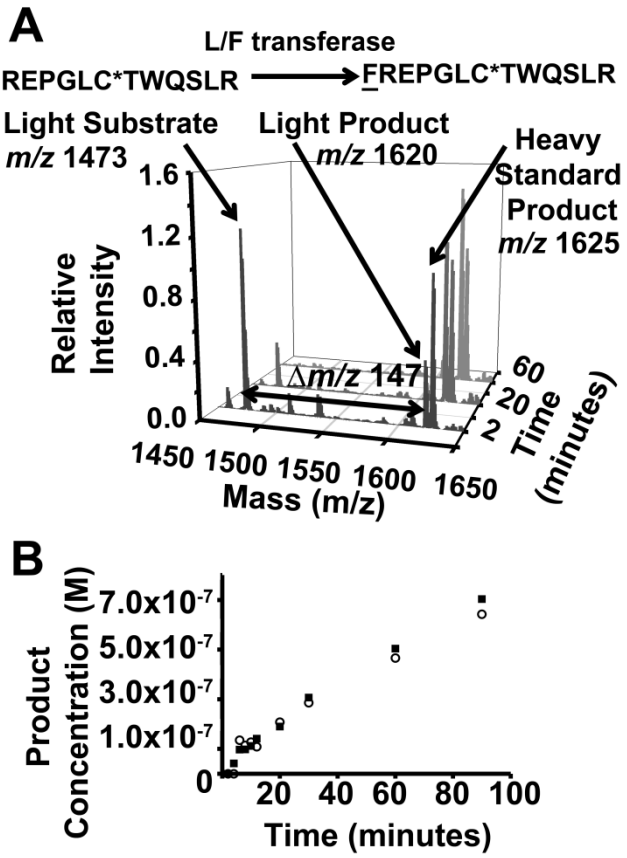

Figure 3-2: **Measuring product formation by mass spectrometry. A)** MALDI-ToF mass spectra of sequential time points of a L/F transferase reaction containing light substrate peptide (REPGLC*TWQSLR, *m/z* 1473). Note the disappearance of light substrate (*m/z* 1473) as the reaction proceeds from 2 to 60 minutes. The appearance of product (FREPGLC*TWQSLR, *m/z* 1620) is also observed and is quantified by the ratio intensities of the light product to the heavy standard peptide (*m/z* 1625). The Δ *m/z* of 147 between light substrate peptide and the light product peptide corresponds to the addition of phenylalanine to the N-terminus of the substrate peptide. **B)** Graphical analysis of product formation when either 5.9 μM of heavy (■) or light (○) substrates are used in enzymatic reactions. Heavy substrate (*m/z* 1478) containing reactions are quantified using an internal light standard product peptide (*m/z* 1620), and light substrate (*m/z* 1473) containing reactions are quantified using an internal heavy standard product peptide (*m/z* 1625). Both heavy and light substrates display a similar rate of product formation indicating there is no difference when either a light or heavy substrate peptide is used in the reaction.

averages of both experimental formats. For simplicity and clarity in the text, all discussions only mention the use of light substrate peptide and heavy standard peptide.

3.2.3. Activity of the D186E Mutant.

Here we report apparent kinetic parameters since experimental limitations prevent the determination of actual K_M and k_{cat} values. The enzymatic reaction involves two substrates and technical limitations prevent the use of experimental conditions where substrates are in sufficient excess. Ideally for the analysis of an enzymatic reaction involving two substrates, one substrate must be maintained in a large excess while the other is varied. For our analysis using varied substrate peptide concentrations, it was difficult to maintain a large excess of the aa-tRNA substrate due to their instability. To obtain sufficient rates of product formation, the amount of enzyme required with respect to substrate concentration is higher than can appropriately be used for the standard quasi-steady state assumption necessary for analysis using Michaelis-Menten models (Laidler 1955, Schnell and Maini 2000). Nevertheless, comparison of the apparent kinetic parameters of the wild-type and mutant L/F transferase enzymes under identical conditions is informative regarding the relative catalytic rates and substrate binding.

Enzymatic reactions of the wild-type and the three D186 L/F transferase mutants (D186A, D186E, and D186N) were performed under standard conditions (3.8 μM enzyme). The results in **Figure 3-3A** (in the

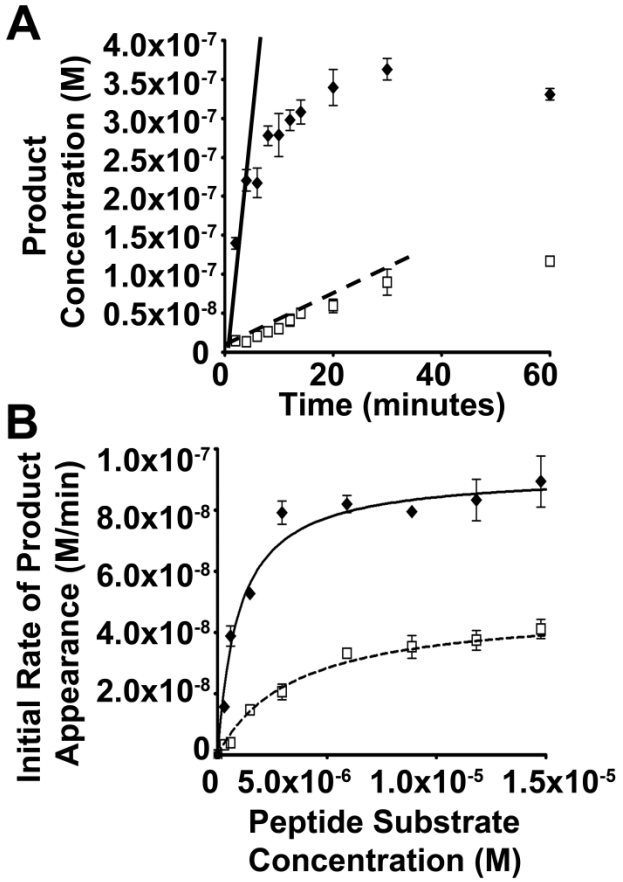

Figure 3-3: Quantifying L/F Transferase activity. **A)** Graphical analysis of the data for the formation of the light product peptide (FREPGLC*TWQSLR, 1620 *m/z*) over time. The data is for (♦) wild-type and the (□) D186E mutant L/F transferase enzymes when using an initial peptide substrate concentration of 0.30 μM. Errors represent duplicate measurements of a single independent experiment. **B)** Analysis of the observed initial reaction rates at different initial substrate concentrations for (♦) wild-type and (□) D186E mutant enzymes. Errors represented are the standard deviation of three independent experiments.

presence of 0.30 μM substrate peptide) clearly demonstrate activity of the D186E point mutant compared to the wild-type. Meanwhile both the D186A and D186N mutants exhibit little or no observable enzymatic activity when tested under identical conditions (data not shown).

A measurable activity was observed for the D186E mutant, its activity was further characterized to compare with the wild-type enzyme. Reactions were performed at peptide substrate concentrations ranging from 0.30 to 14.8 μM and the initial reaction rates determined. The initial reaction rates at different peptide substrate concentrations are plotted in **Figure 3-3B** and numerical values summarized in **Table 3-1**. The apparent kinetic parameters determined are summarized in **Table 3-2**. The data reveals a 2-fold decrease in the apparent k_{cat} and a 3-to-4 fold increase in the apparent K_M (1.0 μM to 3.5 μM) for the D186E mutant when compared to the wild-type. This data for the D186E mutation in conjunction with the apparent inactivity of D186N and D186A suggests that D186 can participate in both substrate binding and catalysis.

3.2.4. Product formation from D186A and D186N Mutants.

The initial apparent inactivity of the D186N and D186A mutations agrees with the proposed protein-based model where D186 is involved in catalysis as previously described (Watanabe *et al.* 2007). However, there are still other possible explanations. Examination of the X-ray crystal structure (**Figure 3-1B**) reveals that D186 may position Q188 for hydrogen bonding to the N-terminus of the substrate peptide and thus facilitate the

Table 3-1: Initial reaction rates of L/F transferase catalyzed peptide bond formation by active site mutants.

L/F transferase	Substrate Peptide Conc. (μM)	Initial Reaction Rate (μM min^{-1})
Wild-type	0.30	0.016 ± 0.002
	0.59	0.039 ± 0.006
	1.5	0.053 ± 0.002
	3.0	0.079 ± 0.006
	5.9	0.082 ± 0.005
	8.9	0.08 ± 0.02
	11.8	0.08 ± 0.01
	14.8	0.09 ± 0.02
D186E	0.30	0.0033 ± 0.0005
	0.59	0.0040 ± 0.001
	1.5	0.015 ± 0.001
	3.0	0.021 ± 0.005
	5.9	0.033 ± 0.003
	8.9	0.035 ± 0.007
	11.8	0.037 ± 0.007
	14.8	0.041 ± 0.006
W49A	0.30	0.0004 ± 0.0004
	0.59	0.0013 ± 0.0001
	1.5	0.0032 ± 0.0002
	3.0	0.0068 ± 0.0006
	5.9	0.0109 ± 0.0005
	8.9	0.0158 ± 0.0003
	11.8	0.019 ± 0.001
	14.8	0.0198 ± 0.0001
M144A	0.30	0.0016 ± 0.0003
	0.59	0.004 ± 0.001
	1.5	0.008 ± 0.001
	3.0	0.011 ± 0.001
	5.9	0.014 ± 0.001
	8.9	0.014 ± 0.002
	11.8	0.014 ± 0.003
	14.8	0.013 ± 0.001
Q188A	0.30	0.0011 ± 0.0006
	0.59	0.0022 ± 0.0005
	1.5	0.0045 ± 0.0008

	3.0	0.003 ± 0.001
	5.9	0.006 ± 0.001
	8.9	0.0067 ± 0.0009
	11.8	0.0055 ± 0.0004
	14.8	0.009 ± 0.001

Errors represented are the standard deviation of three independent experiments (n=3).

Table 3-2: Kinetic parameters of L/F transferase catalyzed peptide bond formation by active site mutants.

L/F transferase	Apparent K_M (peptide) (μM)	Apparent k_{cat} (min^{-1})	Catalytic Efficiency (k_{cat}^{mut}/K_M^{mut} / k_{cat}^{WT}/K_M^{WT})
Wild-type	1.0 ± 0.2	0.0250 ± 0.0008	1.0
D186E	3.5 ± 0.7	0.0129 ± 0.0007	0.147
W49A	16 ± 2	0.011 ± 0.001	0.028
M144A	1.4 ± 0.3	0.0042 ± 0.0002	0.120
Q188A	2.1 ± 0.8	0.0021 ± 0.0002	0.040

Errors represented are the standard deviation of three independent experiments (n=3).

catalysis. For this positioning function of D186, it is predicted that there would be some residual enzymatic activity with the D186N and D186A mutations, while the previously proposed mechanism predicts complete inactivity of these mutants (Watanabe *et al.* 2007).

To further investigate whether the D186A and D186N mutants were truly inactive, we re-evaluated the enzymatic reactions using conditions that would support the detection of product formation even with greatly reduced enzymatic activities. As shown in **Figure 3-4**, reactions using 15 μM L/F transferase revealed product formation for both the D186A and D186N mutants when compared to no enzyme controls. These mass spectra were normalized to the heavy standard product (*m/z* 1625) that is present at a concentration of 1.5 μM. The mass spectra clearly shows the appearance of a light product peptide (*m/z* 1620) at 60 minutes that can be quantified for both the D186A (~12% relative to the standard peptide) and D186N (~7% relative to the standard peptide) L/F transferase mutants **(Figure 3-4B)**. This data suggests that D186A and D186N mutants are both capable of supporting non-ribosomal peptide bond formation.

We were unable to obtain reaction rates for the D186A and D186N mutants at high substrate concentrations. Rates could be measured at substrate concentrations below 3 μM, but at higher concentrations of substrate the combination of high protein and high peptides resulted in such severe noise during MS analysis that product formation could not be reliably quantified.

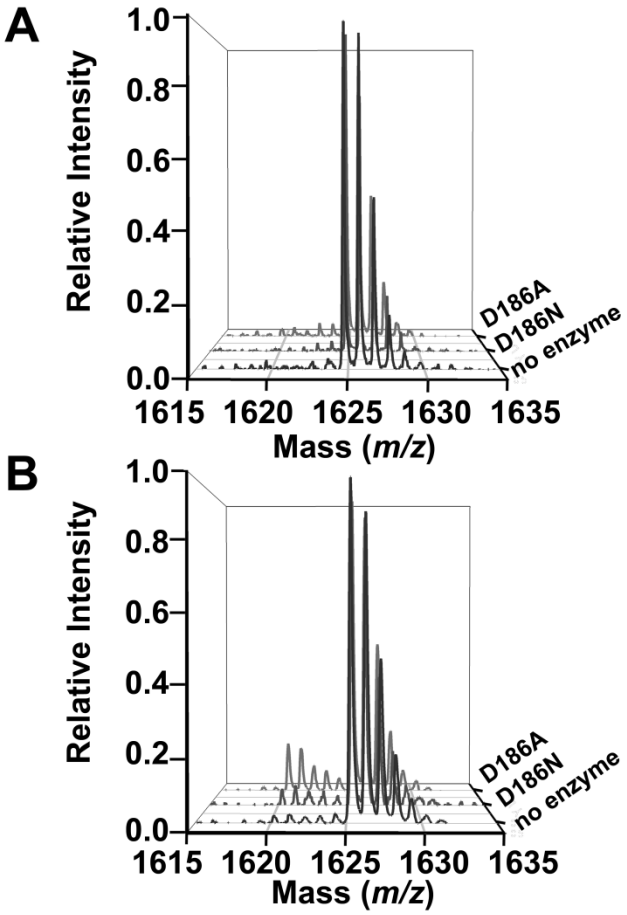

Figure 3-4: Enzymatic activity of D186A and D186N L/F transferase mutants. MALDI-ToF mass spectra comparing L/F transferase reactions containing high amounts (15 μM) of enzyme to a no enzyme control. The mass spectra are aligned in the following order from background to foreground: D186A, D186N and no enzyme control. **A)** After two minutes, no detectable formation of the product peptide (FREPGLC*TWQSLR , *m/z* at 1620) is observed for either D186A or D186N mutants. **B)** After 60 minutes a quantifiable amount of product peptide (1620 *m/z*) is observed in the D186A and D186N reactions when compared to the no enzyme control.

The rates of product formation at low substrate concentrations were measured. Under the single-turnover conditions used (15 μM enzyme and 1.5 μM peptide substrate), initial reaction rates of 0.0017 ± 0.0002 and 0.0006 ± 0.0001 μM min^{-1} were measured for the D186A and D186N mutants respectively. For comparison, these rates are approximately two orders of magnitude slower than the rates observed for the wild-type enzyme at the same concentration (15 μM).

3.2.5. Additional L/F transferase Active Site Mutations.

For additional investigations of L/F transferase activity, four additional point mutants were investigated (Q188A, W49A, W111A, and M144A). All of these mutations have been previously reported to exhibit reduced activity or be completely inactive (Suto *et al.* 2006, Watanabe *et al.* 2007). Q188 is the central amino acid in the proposed electron-relay protein-based mechanism, while W49 and W111 have been proposed to be involved in tRNA binding (Suto *et al.* 2006, Watanabe *et al.* 2007). Q188, W49, and W111 also show significant changes in their binding positions when comparing the substrate analogue bound and product bound complex structures **(Figure 3-1)**, and this suggests that these residues are important for the dynamics of L/F transferase. M144 lines the C-shaped hydrophobic amino acid binding pocket (d_1) for the aa-tRNA substrate (Suto *et al.* 2006, Taki *et al.* 2008). The reduction in activity in W49A, W111A, and M144A mutants have been previously rationalized to be due to its role in recognition and binding of the aa-tRNA substrate

(Suto *et al.* 2006). The reinvestigation of these mutations, with the MALDI-ToF MS method, will validate whether these other mutations are truly inactive and provide a benchmark for comparison to our D186 data.

Under our standard reaction conditions (3.8 μM enzyme), the Q188A, W49A, and M144A mutants resulted in measurable product formation (**Figure 3-5A**) while no product was observed for the W111A mutant. The initial reaction rates at different peptide substrate concentrations are plotted in **Figure 3-5B** and numerical values summarized in **Table 3-1.** The apparent kinetic parameters determined are summarized in **Table 3-2**. The relative rates of the Q188A, W49A, and M144A mutants using the quantitative MALDI-ToF MS method agree with the relative rates derived from conventional radiography methods (Suto *et al.* 2006, Watanabe *et al.* 2007).

To re-evaluate whether W111A is truly inactive, like the D186A and D186N mutations, we examined W111A under high enzyme concentrations (15 μM) identical to the D186A and D186N treatments. **Figure 3-6B** shows the appearance of a product peak (*m/z* 1620) for W111A at 60 minutes that is quantified to be 5% relative to the standard peptide. This data suggests that W111A is also capable of supporting non-ribosomal peptide bond formation.

Figure 3-6C summarizes all substrate binding site mutant investigated where the relative phenylalanine incorporation at 60 minutes

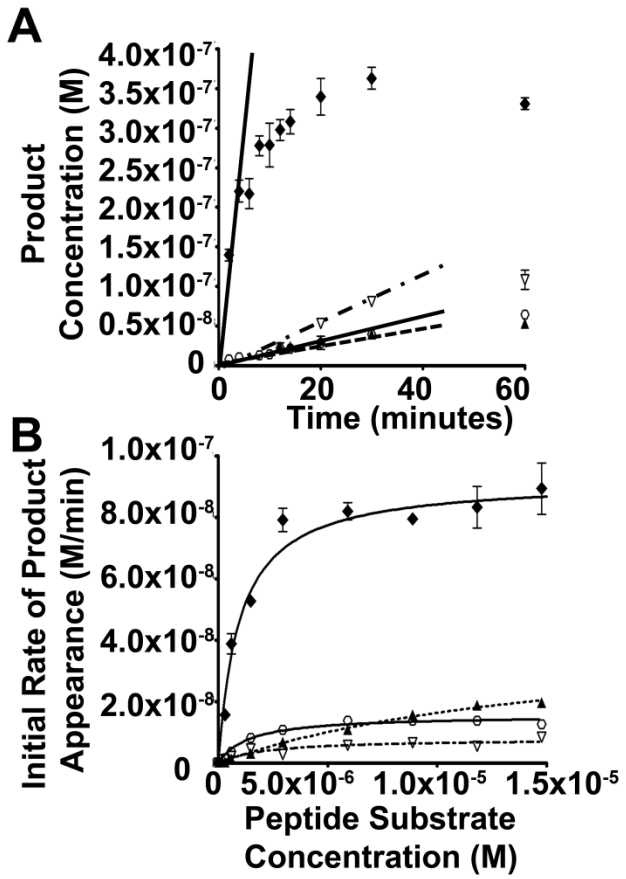

Figure 3-5: Reaction rates of additional L/F transferase mutants activity. A) Analysis of the data for the formation of the light product peptide (FREPGLC*TWQSLR, 1620 *m/z*) over time for (♦) wild-type, (▽) Q188A, (○) M144A, and (▲) W49A enzymes when using an initial substrate concentration of 0.30 μM. Errors represented are a duplicate measurement of a single independent experiment. **B)** Graphic analysis of the observed initial reaction rates at different initial substrate concentrations for (♦) wild-type, (▽) Q188A, (○) M144A, and (▲) W49A enzymes. Errors represented are the standard deviation of three independent experiments.

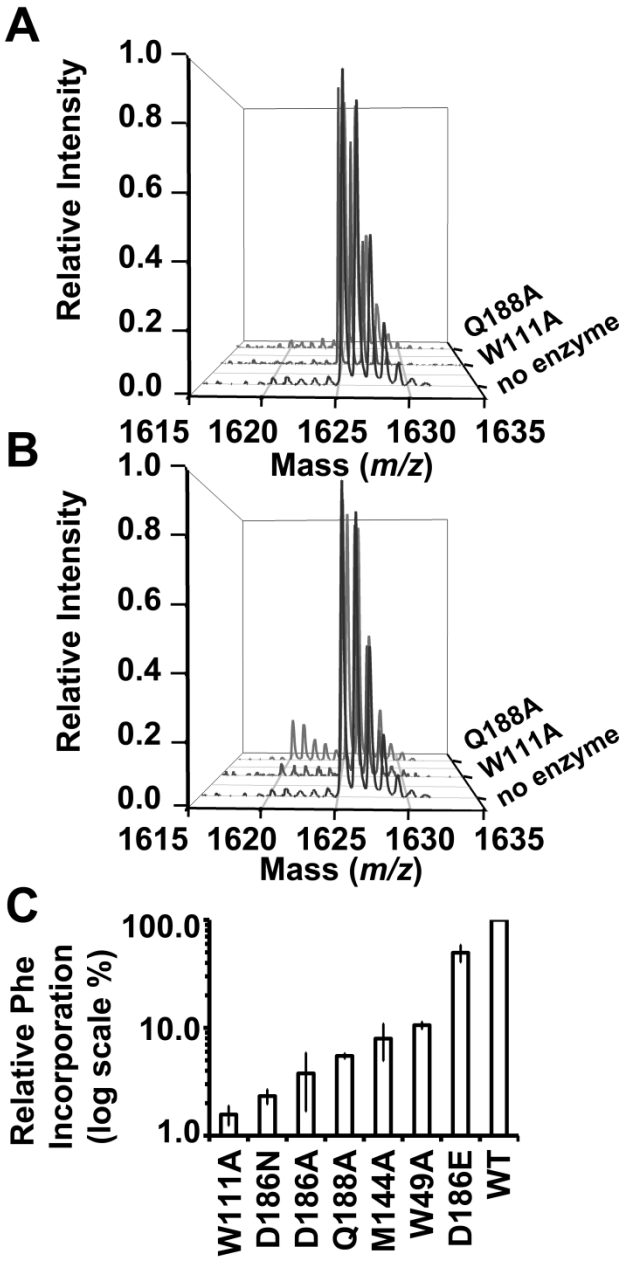

Figure 3-6. Enzymatic activity of W111A and Q188A L/F transferase mutants. MALDI-ToF mass spectra comparing L/F transferase reactions containing high amounts (15 μM) of L/F transferase enzyme and no enzyme control. The mass spectra are aligned in the following order from background to foreground: Q188A, W111A and no enzyme control. **A)** After two minutes, no detectable formation of the product peptide (FREPGLC* TWQSLR, *m/z* at 1620) is observed for either Q188A or W111A. **B)** After 60 minutes the product peptide (1620 *m/z*) is observed in the Q188A and W111A containing reactions in comparison to the no enzyme control. **C)** Phenylalanine incorporation after 60 minutes by various L/F transferase mutants relative to wild-type enzyme in the presence of 1.5 μM substrate peptide and high enzyme concentrations (15 μM). The error bars is represented by the standard deviation of three independent experiments.

in the presence of 1.5 μM substrate peptide and high amounts (15 μM) of L/F transferase is shown. For all mutants, the light product (*m/z* 1620) formed is quantified using the heavy standard product (*m/z* 1625), then normalized to wild-type phenylalanine incorporation. Our data shows that the D186A, D186N, and W111A mutants were capable of catalysis although with rates reduced by two orders of magnitude compared to wild-type. The previously shown reduced activity of Q188A, W49A, and M144A mutants have rates reduced by an order of magnitude compared to wild-type under our reaction conditions. The D186E, although having a 2-fold decrease in phenylalanine incorporation, is the most active mutant in this study.

3.3. Discussion

3.3.1. A Proposed Structural Model for the Function of D186.

As opposed to the previously proposed protein-based, electron-relay model involving both Q188 and D186, we hypothesize an alternative model for the role of D186 and Q188 in L/F transferase catalysis supported by our experimental results. Investigation of the X-ray crystal structure of L/F transferase with a bound product peptide (**Figure 3-1B**) reveals a potential network of hydrogen bonds between Q188, the N-terminus, and side-chain of the arginine of the peptide product. The carboxylic acid side chain of D186 is a hydrogen bond acceptor for the amide side chain of Q188. We propose that D186 facilitates the optimal positioning of Q188 to act as a hydrogen bond acceptor for the main chain

amide and side chain of the arginine peptide substrate (**Figure 3-7A**). This simply positions the N-terminus of the peptide substrate which can then perform a nucleophilic attack on the esterified amino acid of the aa-tRNA.

We propose that the data from the three D186 mutations may be explained as an effect on peptide substrate binding as summarized in **Figure 3-7**. The block arrows represent a potential change in Q188 positioning, when compared to the wild-type, due to the structural changes of D186 mutations. The D186E mutation can still form hydrogen bond Q188 (**Figure 3-7B**), but the longer side-chain results in a less optimal orientation of the peptide substrate by Q188 hence reducing the reaction rate. The side-chain of the alanine in the D186A mutation is smaller and cannot hydrogen bond to Q188 (**Figure 3-7C**). This could result in a significantly altered Q188 placement and subsequent non-productive binding of the peptide substrate. In addition, the D186A mutation may stabilize the Q188 - E156 hydrogen bond observed in the L/F transferase structure in the absence of a bound peptide (**Figure 3-1A**). On the other hand, the D186N mutation could have an even more severe effect on catalysis than the D186A mutation. The D186N mutation reverses the hydrogen bonding character of the Q188 amino acid side chain. D186 becomes a hydrogen bond donor, which could result in the rotation of Q188 to accept the hydrogen bond. The D186N mutation may support an alternative conformation of the Q188 amino acid (**Figure 3-7D**), which

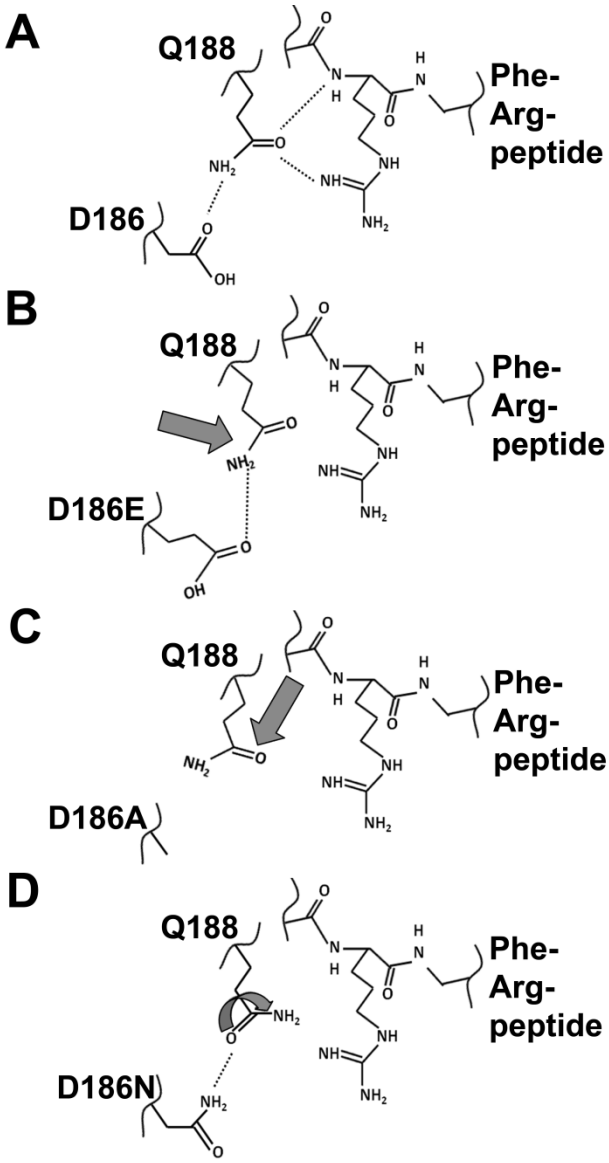

Figure 3-7: Changes in the positioning of Q188 as a consequence of mutations to D186. A) A schematic of the potential hydrogen bonding interactions between D186, Q188 and the product peptide (and substrate peptide by analogy). Grey block arrows in **B)**, **C)**, and **D)** represent a potential change in Q188 positioning as a result of the indicated D186 mutations. **B)** The longer side chain as a result of the D186E mutation can shift the position of Q188 to result in a less ideal positioning of Q188. **C)** The D186A mutation is unable to hydrogen bond to Q188 which may result in a non-productive positioning of Q188. **D)** The D186N mutation may result in the inversion of the Q188 side chain to maintain hydrogen bonding, which would then prevent Q188 from hydrogen bonding to the peptide substrate or product.

would result in the rotation of the amide side chain of Q188 and the amide is being directed towards the N-terminal amine and arginine side chain of the substrate peptide. This is predicted to interfere with substrate binding.

We have attempted to directly measure peptide binding to L/F transferase using both fluorescent and isothermal titration calorimetry methods. In both cases, binding could not be detected (data not shown). This could potentially be the result of a requisite substrate binding order where an aa-tRNA must bind first. This is also in agreement with failed attempts to obtain a crystal structure of L/F transferase with only a bound substrate peptide (Watanabe *et al.* 2007).

3.3.2. Peptide Bond Formation.

If the D186 and Q188 residues are involved in the binding and positioning of the substrate, then what is the catalytic mechanism for L/F transferase? Without evidence to the contrary, we propose that the mechanism of peptide bond formation by L/F transferase is likely to be similar to that proposed for the ribosome (**Figure 3-8**). A current view of the ribosomal peptide bond formation involves the participation of 2'-OH of the peptidyl-tRNA in a mechanism termed 'substrate-assisted catalysis' (Weinger *et al.* 2004, Beringer and Rodnina 2007). This view suggests that the 2'-OH functional group in the peptidyl-tRNA substrate contributes to catalysis by acting as a general acid/base and the ribosome has a more passive role in the specific binding and positioning of the substrate. However, there may be some participation in peptide bond formation

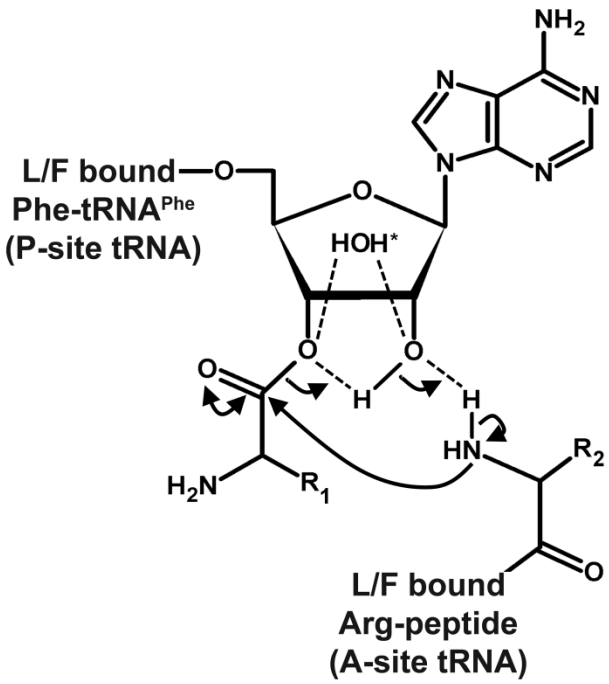

Figure 3-8: Proposed protein-catalyzed peptide bond formation mechanism by L/F transferase. The mechanism of peptide bond formation catalyzed by L/F transferase is proposed to be similar to that for the ribosome. The current view of ribosomal peptide bond formation involves the 2'-OH group of the peptidyl tRNA in a mechanism known as 'substrate assisted' proton shuttling mechanism (Beringer and Rodnina 2007). The L/F transferase bound Phe-tRNA^{Phe} is analogous to the P-site tRNA meanwhile the L/F transferase bound N-terminal Arginine peptide is analogous to the A-site tRNA. The nucleophilic attack of α-amino group onto the ester carbonyl carbon is shown.

chemistry by the ribosome as summarized in a recent review describing the current view of ribosome catalyzed peptide bond formation (Erlacher and Polacek 2008). In the case of L/F transferase, we propose that the 2'-OH of the aa-tRNA substrate would perform an analogous 'substrate assisted catalysis' role. This role of Q188 and D186 in substrate binding and orientation agrees with the data as mutations to these amino acids would alter substrate positioning and reduce observed reaction rates but not entirely prevent catalysis.

With the ribosome only minimally influencing the chemistry of peptide bond formation, it would be surprising for aa-transferases to exhibit a unique protein-based mechanism for peptide bond formation, especially when the reaction rates observed for the L/F transferase enzyme are orders of magnitude slower than the overall rate observed with ribosomes (Erlacher and Polacek 2008). There is no requirement for an increased rate enhancement over that of ribosomes. Nonetheless the relatively large reductions in L/F transferase enzymatic activity as a result of the different mutations investigated demonstrate the significant contribution to catalysis from substrate positioning. Additionally, the recent available biochemical and structural data for the structural homologous $FemX_{Wv}$ alanyl transferase suggests a similar adjacent hydroxyl based proton shuttling mechanism (Fonvielle *et al.* 2010, Fonvielle *et al.* 2013). The tRNA-dependent non-ribosomal peptide bond formation catalytic

mechanisms by L/F transferase and $FemX_{Wv}$ are more similar to the mechanism proposed for the ribosome than previously believed.

3.3.3. Concluding Remarks

By investigating a series of point mutations of both substrates binding sites of L/F transferase enzyme we have determined that despite significant contributions to catalysis by D186 and Q188, these amino acids do not directly participate in the chemistry of peptide bond formation. Our data supports a model where D186 is involved in the positioning of Q188, and Q188 contributes to the optimal positioning of the peptide substrate for non-ribosomal peptide bond formation. Overall we propose that L/F transferase does not directly participate in the chemistry of peptide bond formation but catalyzes the reaction by binding and orientating the substrates for reaction via a mechanism that has been described for ribosomes.

3.4. References

Abramochkin, G., and Shrader, T.E. (1995) The leucyl/phenylalanyl-tRNA-protein transferase. Overexpression and characterization of substrate recognition, domain structure, and secondary structure. *J.Biol.Chem.* **270,** 20621-20628

Beringer, M., and Rodnina, M.V. (2007) The ribosomal peptidyl transferase. *Mol.Cell.* **26,** 311-321

Desiderio, D.M., and Kai, M. (1983) Preparation of stable isotope-incorporated peptide internal standards for field desorption mass spectrometry quantification of peptides in biologic tissue. *Biomed.Mass Spectrom.* **10,** 471-479

Dong, X., Kato-Murayama, M., Muramatsu, T., Mori, H., Shirouzu, M., Bessho, Y., and Yokoyama, S. (2007) The crystal structure of leucyl/phenylalanyl-tRNA-protein transferase from *Escherichia coli. Protein Sci.* **16,** 528-534

Ebhardt, H.A., Xu, Z., Fung, A.W., and Fahlman, R.P. (2009) Quantification of the post-translational addition of amino acids to proteins by MALDI-TOF mass spectrometry*. Anal.Chem.* **81,** 1937-1943

Erlacher, M.D., and Polacek, N. (2008) Ribosomal catalysis: the evolution of mechanistic concepts for peptide bond formation and peptidyl-tRNA hydrolysis*. RNA Biol.* **5,** 5-12

Fonvielle, M., Chemama, M., Lecerf, M., Villet, R., Busca, P., Bouhss, A., Etheve-Quelquejeu, M., and Arthur, M. (2010) Decoding the logic of the tRNA regiospecificity of nonribosomal FemX(Wv) aminoacyl transferase*. Angew.Chem.Int.Ed Engl.* **49,** 5115-5119

Fonvielle, M., Li de La Sierra-Gallay, I., El-Sagheer, A.H., Lecerf, M., Patin, D., Mellal, D., Mayer, C., Blanot, D., Gale, N., Brown, T., van Tilbeurgh, H., Etheve-Quelquejeu, M., and Arthur, M. (2013) The structure of FemX(Wv) in complex with a peptidyl-RNA conjugate: mechanism of aminoacyl transfer from Ala-tRNA(Ala) to peptidoglycan precursors*. Angew.Chem.Int.Ed Engl.* **52,** 7278-7281

Fung, A.W., Ebhardt, H.A., Abeysundara, H., Moore, J., Xu, Z., and Fahlman, R.P. (2011) An alternative mechanism for the catalysis of peptide bond formation by L/F transferase: substrate binding and orientation*. J.Mol.Biol.* **409,** 617-629

Ichetovkin, I.E., Abramochkin, G., and Shrader, T.E. (1997) Substrate recognition by the leucyl/phenylalanyl-tRNA-protein transferase. Conservation within the enzyme family and localization to the trypsin-resistant domain*. J.Biol.Chem.* **272,** 33009-33014

Kaji, A., Kaji, H., and Novelli, G.D. (1965a) Soluble Amino Acid-Incorporating System. I. Preparation of the System and Nature of the Reaction*. J.Biol.Chem.* **240,** 1185-1191

Kaji, A., Kaji, H., and Novelli, G.D. (1965b) Soluble Amino Acid-Incorporating System. II. Soluble Nature of the System and the Characterization of the Radioactive Product*. J.Biol.Chem.* **240,** 1192-1197

Kuno, A., Taki, M., Kaneko, S., Taira, K., and Hasegawa, T. (2003) Leucyl/phenylalanyl (L/F)-tRNA-protein transferase-mediated N-terminal

specific labelling of a protein in vitro. *Nucleic Acids Res.Suppl.* **(3),** 259-260

Laidler, K. (1955) Theory of the Transient Phase in Kinetics, with Special Reference to Enzyme Systems. *Canadian Journal of Chemistry-Revue Canadienne De Chimie.* **33,** 1614-1624

Schnell, S., and Maini, P.K. (2000) Enzyme kinetics at high enzyme concentration. *Bull.Math.Biol.* **62,** 483-499

Suto, K., Shimizu, Y., Watanabe, K., Ueda, T., Fukai, S., Nureki, O., and Tomita, K. (2006) Crystal structures of leucyl/phenylalanyl-tRNA-protein transferase and its complex with an aminoacyl-tRNA analog. *EMBO J.* **25,** 5942-5950

Taki, M., Kuroiwa, H., and Sisido, M. (2008) Chemoenzymatic transfer of fluorescent non-natural amino acids to the N terminus of a protein/peptide. *Chembiochem.* **9,** 719-722

Wagner, A.M., Fegley, M.W., Warner, J.B., Grindley, C.L., Marotta, N.P., and Petersson, E.J. (2011) N-terminal protein modification using simple aminoacyl transferase substrates. *J.Am.Chem.Soc.* **133,** 15139-15147

Watanabe, K., Toh, Y., Suto, K., Shimizu, Y., Oka, N., Wada, T., and Tomita, K. (2007) Protein-based peptide-bond formation by aminoacyl-tRNA protein transferase. *Nature.* **449,** 867-871

Weinger, J.S., Parnell, K.M., Dorner, S., Green, R., and Strobel, S.A. (2004) Substrate-assisted catalysis of peptide bond formation by the ribosome. *Nat.Struct.Mol.Biol.* **11,** 1101-1106

Weinger, J.S., and Strobel, S.A. (2006) Participation of the tRNA A76 hydroxyl groups throughout translation. *Biochemistry.* **45,** 5939-5948

Chapter 4

Probing tRNA Binding to L/F transferase with Substrate Analogues

A version of this chapter is published in:

Fung AW, Ebhardt HA, Krishnakumar KS, Moore J, Xu Z, Strazewski P, and Fahlman RP (2014) Probing the Leucyl/Phenylalanyl tRNA Protein Transferase Active Site with tRNA Substrate Analogues. *Protein and Peptide Letters*. 21(7): 603-614.

4.1. Introduction

Despite the many discovered biological functions of aa-transferases (see **1.1.**), there is a lack of understanding in the regulation and molecular mechanisms of these enzymes. Currently, there are no known specific chemical inhibitors for these enzymes. The eubacterial L/F transferase has been described to be inhibited by divalent cations (i.e. Mg^{2+}, Ca^{2+}, and Mn^{2+}) (Leibowitz and Soffer 1970). The eukaryotic ATE1 has been demonstrated to be inhibited by cysteine reactive chemicals (Klemperer and Pickart 1989, Berleth *et al.* 1992, Li and Pickart 1995a, Li and Pickart 1995b) as well as through oxidation by heme (Hu *et al.* 2008). The basis of inhibition by these chemical reactions is currently unclear, however, these are non-specific reactions that either oxidize or modify the enzyme. Some recent data also suggests that the inhibition by heme may be indirectly through the inhibition of arginyl-tRNA synthetase (Yang *et al.* 2010). To date, the only structurally targeted inhibition of aa-transferases has been by the application of excess amounts of substrate peptide or use of non-hydrolyzable tRNA substrate analogues to out-compete endogenous substrates (Leibowitz and Soffer 1970, Baker and Varshavsky 1991).

Initial biochemical data suggests that the aa-tRNA recognition by L/F transferase is mainly through the 3' terminal adenosine (A_{76}) and the esterified amino acid (Leibowitz and Soffer 1971, Abramochkin and Shrader 1996). Since these earlier reports, several X-ray crystal

structures of L/F transferase with tRNA substrate analogues bound have become available providing valuable insights into the molecular understanding of the mechanism of tRNA recognition by L/F transferase (see **1.4.2.**) (Suto *et al.* 2006, Dong *et al.* 2007, Watanabe *et al.* 2007).

Although both tRNA substrate analogues bind to the central cleft active site of L/F transferase, closer examination revealed significant differences in their orientation and binding to L/F transferase. The chemical differences between the minimal substrate phenylalanyl adenosine (rA-Phe) and the inert analogue puromycin include an ester versus amide linkage, the presence of dimethylation modification on the adenine base, and the presence of a methoxy-modification on the phenylalanine side chain (**Figure 4-1A**). Each individual chemical difference may play different roles in the binding to L/F transferase. For example, the rigidity of an amide linkage may alter the flexibility and dynamics of the molecule. Given that there are differences in the substrate analogue-bound crystal structures and that there are no crystal structures solved for L/F transferase in complex with an intact aa-tRNA bound, the molecular insights derived from these structures remain within the 3'-end of an aa-tRNA.

Here we utilize published X-ray crystal structures (Suto *et al.* 2006, Dong *et al.* 2007, Watanabe *et al.* 2007) and the quantitative MALDI-ToF mass spectrometry based activity assay (Ebhardt *et al.* 2009, Fung *et al.* 2011) to functionally investigate the binding difference between rA-Phe

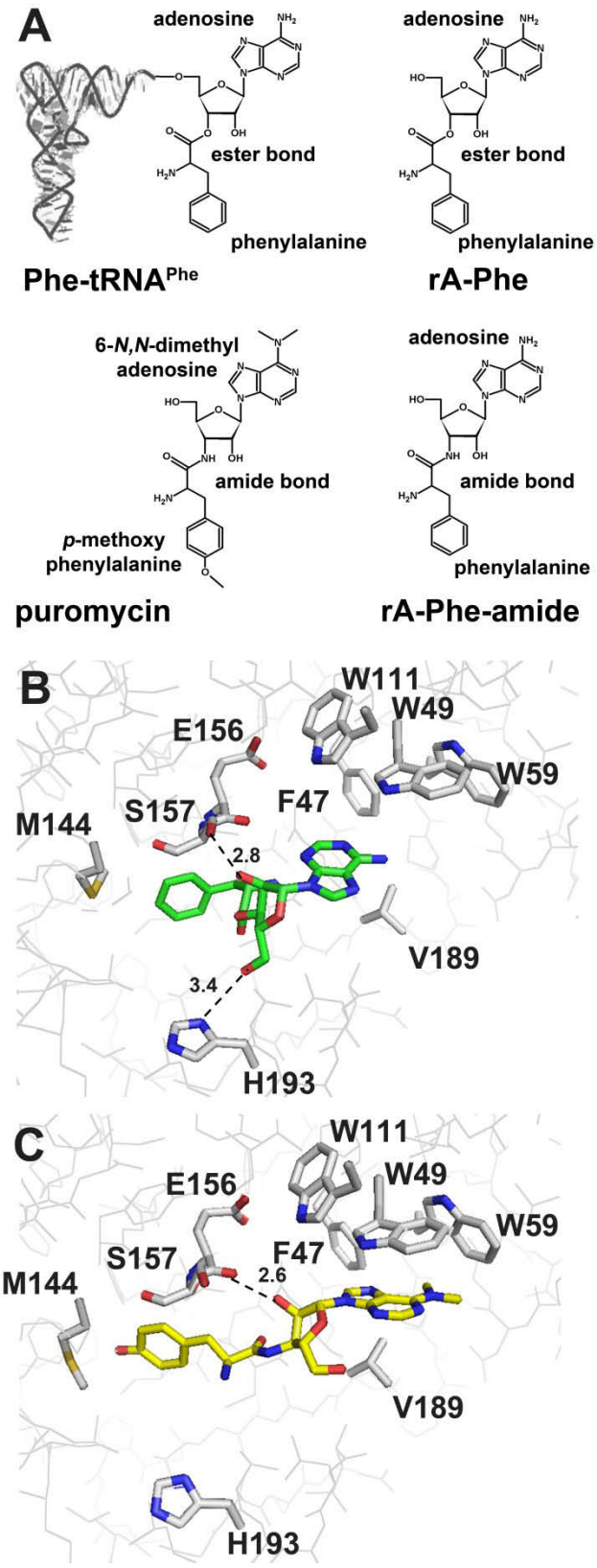

Figure 4-1: Differential binding of tRNA substrate analogues to L/F transferase. A) Chemical structures of phenylalanyl-tRNA^{Phe} (Phe-tRNA^{Phe}), substrate analogues rA-Phe, puromycin and rA-Phe-amide. Puromycin is an antibiotic that mimics the 3' terminus of an aa-tRNA. X-ray crystal structure of L/F transferase in complex with substrate analogue **B)** rA-Phe (green, PDB ID: 2Z3K) or **C)** puromycin (yellow, PDB ID: 2DPT) reveal differences in their binding to L/F transferase. The amino acid moiety of rA-Phe and puromycin both occupy the C-shaped hydrophobic pocket; however the ribose and the adenosine structures are bound differently. Residues that are important for hydrogen bonding, π-π stacking and hydrophobic interactions are labeled. This image was prepared from a Protein Data Bank (PDB) file using PyMOL Molecular Graphics System, version 1.3, Schrödinger, LLC.

and puromycin to L/F transferase. Through measuring their inhibition of L/F transferase catalyzed reactions, we have determined that both substrate analogues bind to L/F transferase with similar affinities. Further analysis of available X-ray crystal structures revealed that the presence of dimethyl-modified adenine base and methoxy-modified phenylalanine have opposite effects on binding, which serendipitously results in a similar binding affinity. The modified adenine base of puromycin has favourable interactions with L/F transferase via π-π stacking and additional hydrophobic interactions, whereas the modified phenylalanine of puromycin hinders binding via a steric clash with M144 in the C-shaped hydrophobic pocket. Through structural analysis, mutagenesis and enzymatic activity assays, we have determined that the inhibition and hence binding of puromycin can be enhanced by increasing the size of the hydrophobic binding pocket in L/F transferase. Future investigations may lead to the developments of improved inhibitors for aa-transferases.

4.2. Results

4.2.1. L/F transferase tRNA Recognition via Substrate Analogue Bound Structures

The chemical structures of the tRNA substrate analogues, rA-Phe and puromycin, observed in the X-ray crystal structures are shown in **Figure 4-1A** (Suto *et al.* 2006, Watanabe *et al.* 2007). rA-Phe is the minimal donor substrate, which is chemically equivalent to the 3' end of an aa-tRNA (Watanabe *et al.* 2007, Wagner *et al.* 2011). However rA-Phe

lacks the remaining of the tRNA body for additional interactions, which results in significantly weaker binding affinity (reported K_M (rA-Phe) = 124 μM versus K_M (Phe-tRNA^{Phe}) = ~2 μM) albeit there is no significant differences in k_{cat} (Abramochkin and Shrader 1996, Wagner *et al.* 2011). The binding and inhibition of *E. coli* L/F transferase by the antibiotic puromycin was documented over 40 years ago (Leibowitz and Soffer 1970). Puromycin (3'-deoxy-*N,N*-dimethyl-3'-[(*O*-methyl-$_L$-tyrosyl) amino] adenosine) is a non-hydrolyzable analogue of the 3' terminus of an aa-tRNA (Abramochkin and Shrader 1996, Suto *et al.* 2006). The *p*-methoxyphenylalanine group of puromycin corresponds to the phenylalanine side chain of the amino acid and the 6-*N, N*-dimethyladenosine group corresponds to the 3' terminal adenosine (A_{76}) of an aa-tRNA. A key chemical difference between puromycin and an aa-tRNA is that puromycin contains a non-hydrolyzable amide linkage to the ribose. Puromycin is widely known for its inhibition of ribosomal protein synthesis (Ennis 1965a, Ennis 1965b) and the mechanism in which puromycin inhibits L/F transferase has been suggested to be a competition for the aa-tRNA substrate binding site (Horinishi *et al.* 1975, Abramochkin and Shrader 1996, Suto *et al.* 2006).

rA-Phe (green, **Figure 4-1B**) and puromycin (yellow, **Figure 4-1C**) both bind to the active site on the central cleft of L/F transferase (Suto *et al.* 2006, Watanabe *et al.* 2007). In both structures, the C-shaped hydrophobic pocket accommodates the amino acid side chain and the

ribose adopts the C_3' endo conformation as a typical RNA ribose does. However the ester bond linkage in rA-Phe and the amide bond linkage in puromycin may alter the geometry and rigidity of the analogues. Closer examination of the complex structures reveals differences in the binding orientations for the ribose and the adenine. In the rA-Phe bound structure, the adenosine appears mobile to undergo a conformational change upon catalysis (Watanabe *et al.* 2007). The ribose group of rA-Phe forms two hydrogen bonds with L/F transferase (2'-OH with S157 and 5'-OH with H193). Conversely, the ribose and 6-*N, N*-dimethyladenosine groups of puromycin are rotated and extended when compared to the rA-Phe bound structure. This rotation and extension seems to optimize the binding of 6-*N,N*-dimethyladenosine to L/F transferase via π-π base stacking interactions with W49 (which is further stabilized with stacking interactions with W111), and additional hydrophobic interaction with F47, W59, and V189. This hydrophobic interaction is not observed in the rA-Phe bound structure. Additionally the 2'-OH of puromycin is hydrogen bonded to E156 instead of S157. Given that there are differences in the binding of tRNA substrate analogues and there is no crystal structure solved for L/F transferase in complex with an intact aa-tRNA, the molecular insights derived from these structures remain somewhat uncertain. Here we utilize published X-ray crystal structures (Suto *et al.* 2006, Dong *et al.* 2007, Watanabe *et al.* 2007) and the quantitative MALDI-ToF mass spectrometry based activity assay (Ebhardt *et al.* 2009, Fung *et al.* 2011)

to functionally investigate the binding differences between rA-Phe and puromycin to L/F transferase.

4.2.2. Binding of substrate analogues to L/F transferase by inhibition assays

To investigate the interactions of substrate analogues to L/F transferase, we used an indirect approach by measuring their inhibition of L/F transferase catalyzed reactions. To generate a non-hydrolyzable analogue of rA-Phe, rA-Phe-amide (3'-deoxy-3'-[($_L$-phenylalanyl) amino] adenosine) was synthesized (**Figure 4-1A**). rA-Phe-amide, like rA-Phe, retains the structure of unmodified adenine base and phenylalanine amino acid side chain but contains the non-hydrolyzable amide linkage. The effects of modified adenine base and modified phenylalanine on binding can now be compared since both puromycin and rA-Phe-amide contain amide bond linkages. Additionally, PheRS is known to aminoacylate at the 2'-OH of A_{76} of $tRNA^{Phe}$ and the aminoacyl moiety then transfers to the 3'-OH via trans-esterification (Eriani *et al.* 1990). Since an amide linkage would lock the aminoacyl moiety at either the 2' or 3' position and Watanabe *et al.* have demonstrated that L/F transferase does not utilize 2'-Phe-$tRNA^{Phe}$, the synthesized rA-Phe-amide resembles the trans-acylated 3'-Phe-$tRNA^{Phe}$ (Watanabe *et al.* 2007).

Many two-substrate enzyme-catalyzed reactions obey the Michaelis-Menten equation when one substrate concentration is fixed in excess while the other is varied. Since puromycin and rA-Phe-amide are

analogues of an aa-tRNA, it would be ideal to measure L/F transferase activity with varied tRNA concentrations. However, there are technical challenges to maintain specific concentrations of aa-tRNA substrate due to their instability since the exact concentration is unknown under continuous aminoacylation. For our analysis, we therefore varied substrate peptide concentrations instead. To obtain sufficient rates of product formation with mutations and inhibition, we have used higher enzyme concentrations than the standard quasi-steady state assumption necessary for Michaelis-Menten models (Laidler 1955, Schnell and Maini 2000). Nonetheless, the observed rates and apparent inhibition parameters of the wild-type and mutant L/F transferase enzymes under identical conditions can be directly compared and are informative regarding their relative catalytic efficiency, substrate binding and inhibition.

We have evaluated the inhibition of wild-type L/F transferase catalyzed reactions by puromycin and rA-Phe-amide by measuring the apparent inhibition constant (K_i). To quantify the activity of the wild-type L/F transferase enzymes in the presence of puromycin or rA-Phe-amide, we applied a modified quantitative MALDI-ToF MS assay (Ebhardt *et al.* 2009) with the alterations as previously described (Fung *et al.* 2011). L/F transferase wild-type activity was measured at different peptide substrate concentrations in the presence of four different concentrations of puromycin or rA-Phe-amide (0, 0.24, 0.48, and 0.95 mM). Initial rates of product formation determined for the wild-type enzyme are listed in **Table**

4-1, while the data is summarized graphically in **Figure 4-2**. The apparent K_i of puromycin for wild-type L/F transferase is calculated to be 425 ± 36 μM (curve-fitting analysis by GraphPad Prism Version 5.02). This is the first reporting of the apparent K_i of puromycin for L/F transferase. Our data is in agreement with previous investigations reporting 70% inhibition at 0.25 mM puromycin (Leibowitz and Soffer 1970) and 75% inhibition at 0.40 mM puromycin (Horinishi *et al.* 1975, Abramochkin and Shrader 1996) under different reaction conditions. A similar analysis was performed with rA-Phe-amide and the apparent K_i of rA-Phe-amide for wild-type L/F transferase is calculated to be 659 ± 72 μM. A comparison of the apparent K_i of puromycin (425 μM) and rA-Phe-amide (659 μM) for wild-type L/F transferase, suggests that the binding affinities of L/F transferase to both substrate analogues is similar. With similar binding affinities, it cannot be determined whether both molecules bind differently with similar affinities as indicated by the X-Ray crystal structures or whether the molecules are simply binding L/F transferase in a similar fashion.

Further analysis of the secondary plots (**Figure 4-3** and **4-4**) and curve-fitting data reveals that puromycin is a non-competitive inhibitor of the enzyme under our reaction conditions. Due to the complex nature of two-substrate reactions, a particular inhibitor that is theoretically competitive in nature may not necessarily result in a characteristic competitive inhibition pattern (Palmer 1991). Non-competitive inhibition is

Table 4-1: **Initial reaction rates of L/F transferase catalyzed peptide bond dormation during puromycin and rA-Phe-amide inhibition.**

A) L/F transferase wild-type and puromycin

L/F transferase Wild Type	Substrate Peptide Conc. (μM)	Initial Reaction Rate (μM min^{-1})
Wild Type 0.0 mM puromycin	0.30	0.068 ± 0.012
	0.59	0.075 ± 0.020
	1.5	0.11 ± 0.03
	3.0	0.12 ± 0.02
	5.9	0.13 ± 0.02
	8.9	0.14 ± 0.02
	11.8	0.14 ± 0.02
	14.8	0.15 ± 0.02
Wild Type 0.24 mM puromycin	0.30	0.022 ± 0.003
	0.59	0.047 ± 0.025
	1.5	0.070 ± 0.025
	3.0	0.078 ± 0.025
	5.9	0.083 ± 0.021
	8.9	0.083 ± 0.017
	11.8	0.088 ± 0.019
	14.8	0.098 ± 0.020
Wild Type 0.48 mM puromycin	0.30	0.031 ± 0.012
	0.59	0.030 ± 0.005
	1.5	0.058 ± 0.022
	3.0	0.053 ± 0.018
	5.9	0.059 ± 0.004
	8.9	0.060 ± 0.007
	11.8	0.069 ± 0.019
	14.8	0.067 ± 0.025
Wild Type 0.95 mM puromycin	0.30	0.015 ± 0.005
	0.59	0.033 ± 0.010
	1.5	0.043 ± 0.013
	3.0	0.042 ± 0.011
	5.9	0.044 ± 0.014
	8.9	0.044 ± 0.010
	11.8	0.046 ± 0.020
	14.8	0.040 ± 0.013

Errors represented are the standard deviation of three independent experiments (n=3).

B) L/F transferase wild-type and rA-Phe-amide

L/F transferase	Substrate Peptide Conc. (μM)	Initial Reaction Rate (μM min^{-1})
Wild Type 0.0 mM rA-Phe-amide	0.59	0.041 ± 0.006
	3.0	0.072 ± 0.015
	5.9	0.081 ± 0.017
	11.8	0.086 ± 0.014
Wild Type 0.24 mM rA-Phe-amide	0.59	0.040 ± 0.007
	3.0	0.050 ± 0.011
	5.9	0.062 ± 0.003
	11.8	0.062 ± 0.005
Wild Type 0.48 mM rA-Phe-amide	0.59	0.029 ± 0.008
	3.0	0.037 ± 0.009
	5.9	0.045 ± 0.011
	11.8	0.050 ± 0.010
Wild Type 0.95 mM rA-Phe-amide	0.59	0.017 ± 0.008
	3.0	0.027 ± 0.006
	5.9	0.035 ± 0.010
	11.8	0.036 ± 0.003

Errors represented are the standard deviation of three independent experiments (n=3).

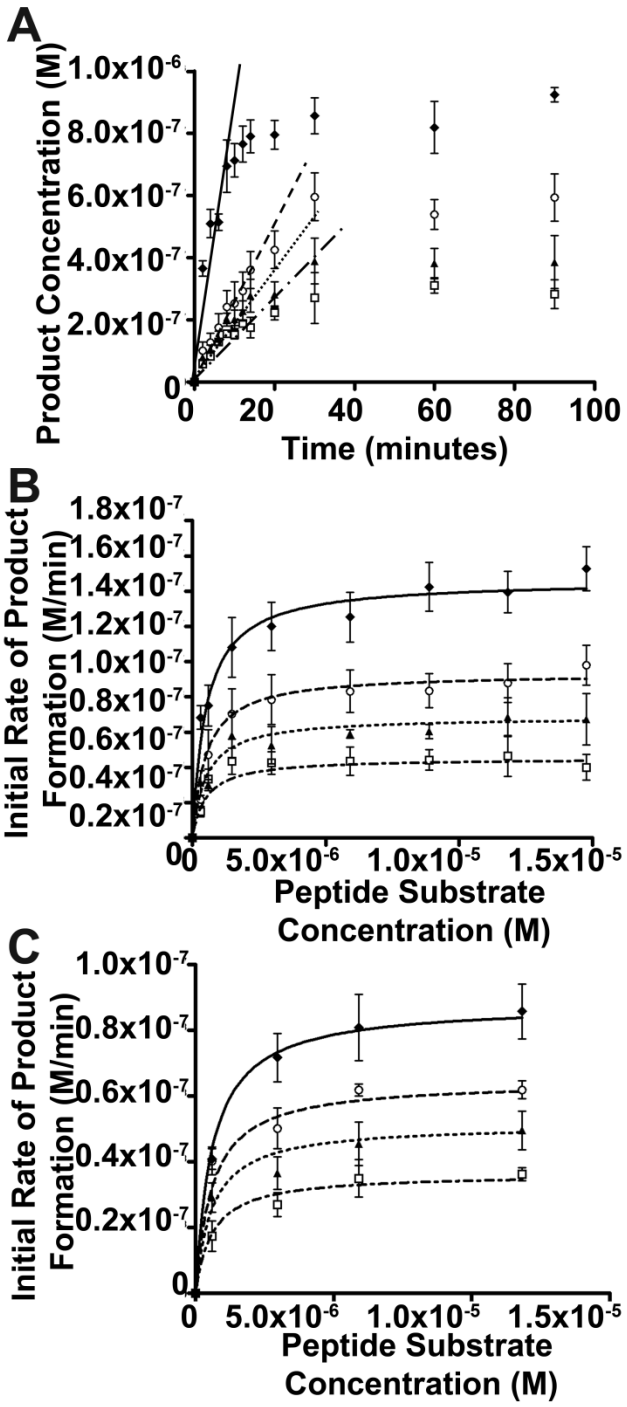

Figure 4-2: Inhibition of wild-type L/F transferase by puromycin and rA-Phe-amide. A) Graphical analysis of product formation over time for wild-type L/F transferase in the presence of 0.0 (♦), 0.24 (○), 0.48 (▲), and 0.95 mM (□) of puromycin when using an initial peptide substrate concentration of 0.30 μM. Errors represent triplicate measurements of a single independent experiment. Initial rate of product formation is calculated from the slope of the linear tangent line drawn to the curve. **B)** and **C)** A graphical display of initial rate of product formation against peptide substrate concentration for wild-type L/F transferase in the presence of 0.0 (♦), 0.24 (○), 0.48 (▲), and 0.95 mM (□) of **B)** puromycin or **C)** rA-Phe-amide. Errors represented are the standard deviation of three independent experiments.

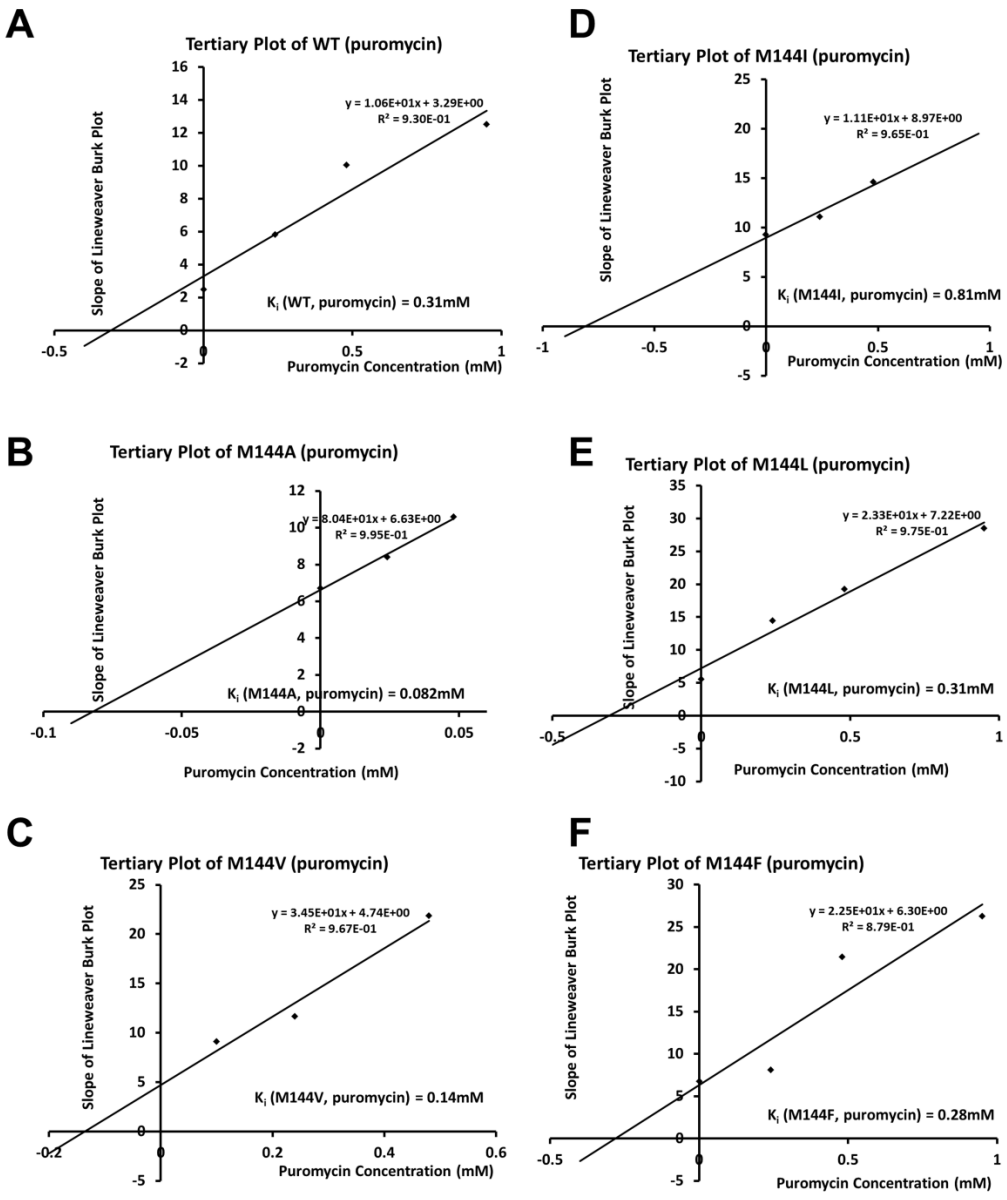

Figure 4-3: Secondary plots of slope of Lineweaver-Burk plot against puromycin concentrations for determining puromycin inhibition constant on A) wild-type L/F transferase, B) M144A, C) M144V, D) M144I, E) M144L, and F) M144F.

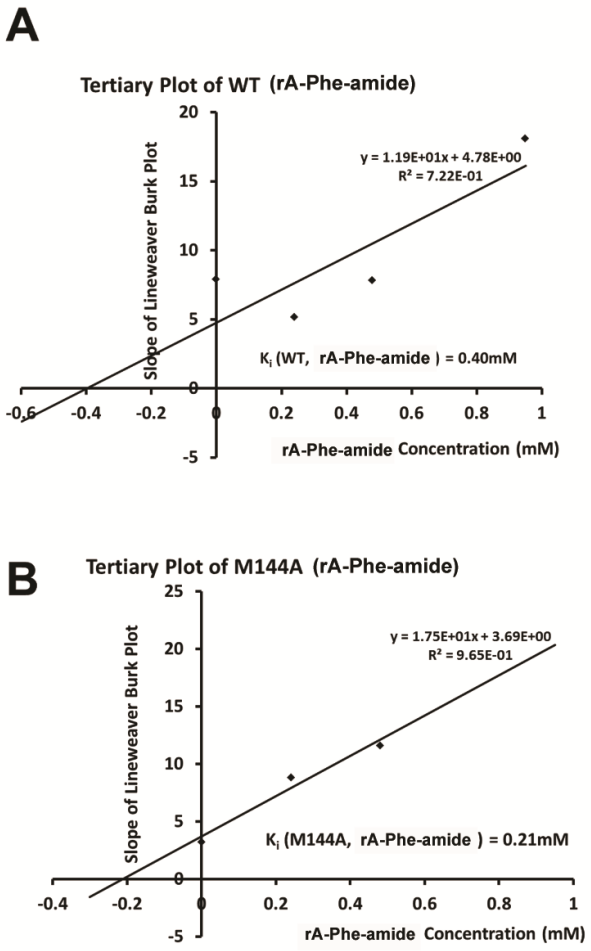

Figure 4-4: **Secondary plots of slope of Lineweaver-Burk plot against puromycin concentrations for determining rA-Phe-amide inhibition constant on A) wild-type L/F transferase and B) M144A.**

characterized by reducing the reaction rate without altering the apparent affinity or K_M for the substrate. In other words, puromycin and the substrate polypeptide probably do not compete for the same binding site on L/F transferase. As puromycin is an analogue of the aa-tRNA substrate and our assay varies peptide substrate concentrations and monitors the appearance of the product peptide, a non-competitive mode of inhibition agrees with the data. Additionally the difference of several orders of magnitude when comparing the K_i of puromycin (425 μM) or rA-Phe-amide (659 μM) with the reported K_M of Phe-tRNA^{Phe} (~2 μM) (Abramochkin and Shrader 1996), despite under different reaction conditions, suggests significantly lower affinity for puromycin or rA-Phe-amide than the natural aa-tRNA substrate.

4.2.3. Steric clash model between M144 and p-methoxyphenylalanine of puromycin

Figure 4-1C shows that the 6-*N, N*-dimethyladenosine group of puromycin is stabilized by π-π stacking interactions with W49, which is further stabilized stacking interactions with W111. Additional hydrophobic contacts with F47, W59, and V189 also further contribute to the interactions with 6-*N, N*-dimethyladenosine group of puromycin. These forces together would greatly favour puromycin binding. However puromycin only gives a modest inhibition of L/F transferase (high μM range), while puromycin is a potent inhibitor of ribosomal synthesis (low μM range) (Starck and Roberts 2002). We hypothesize that the presence

of a modified adenine base and modified phenylalanine may have different and opposite effects on binding, which serendipitously results in similar apparent binding affinity between the two analogues (puromycin and rA-Phe-amide). To examine whether the additional *p*-methoxy group of puromycin influences binding, we further analyzed available X-ray crystal structures with bound substrate analogues or product peptide and their interaction with the C-shaped hydrophobic pocket. It has been suggested that the substrate specificity of L/F transferase selecting an aa-tRNA is through the specific hydrophobic interactions in this pocket (d_1) (Suto *et al.* 2006). One amino acid of interest lining the hydrophobic pocket is methionine 144 (M144).

Figure 4-5A shows the surface representation of L/F transferase with puromycin bound. **Figure 4-5B** shows a superimposed complex structure of L/F transferase with bound puromycin (yellow) (Suto *et al.* 2006), bound substrate analogue rA-Phe (green) (Watanabe *et al.* 2007), bound product peptide (cyan) (Watanabe *et al.* 2007), and no substrate bound (pink) (Dong *et al.* 2007). The superimposed structure reveals that the M144 side chain exists in multiple conformations depending on which molecule is bound to the protein. Interestingly, M144 is rotated and pointed away from the hydrophobic amino acid pocket only in the puromycin-bound state (yellow), but not in the other states (green, cyan and pink).

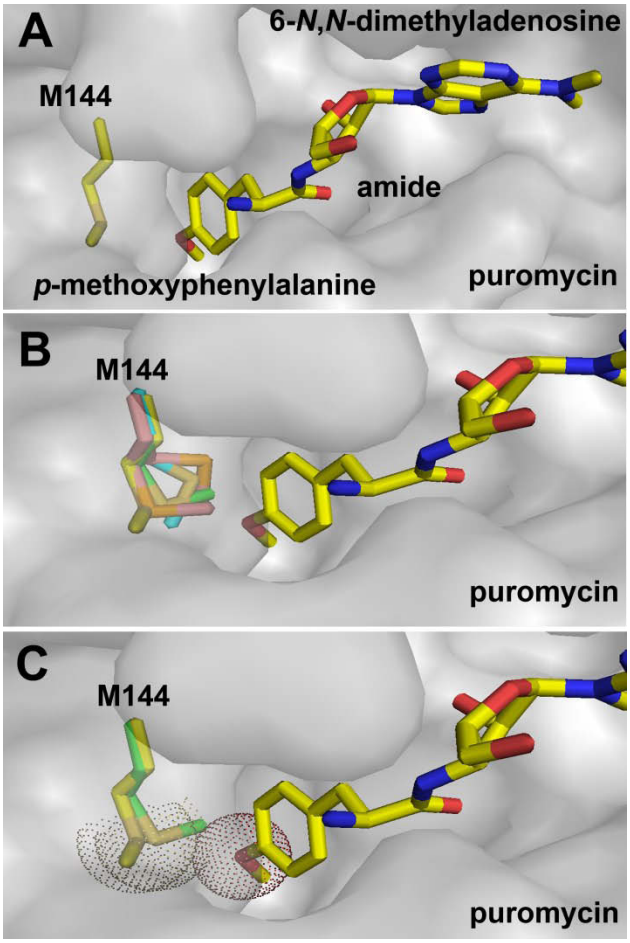

Figure 4-5: X-ray crystal structures of L/F transferase with puromycin binding. A) X-ray crystal structure of puromycin-bound L/F transferase wild-type. This image was prepared from a Protein Data Bank (PDB) file, PDB 2DPT, using PyMOL Molecular Graphics System, version 1.3, Schrödinger, LLC. **B)** A superimposed X-ray crystal structure of L/F transferase demonstrates the multiple conformations of M144. The side chain of M144 with puromycin bound (yellow, PDB ID: 2DPT), substrate analogue adenosine phenylalanine bound (rA-Phe, green, PDB ID: 2Z3K), product peptide bound (cyan, PDB ID: 2Z3L), and with nothing bound (pink, PDB ID: 2CXA) are shown. Specifically notice that M144 rotates and points away from the pocket when puromycin is bound (yellow) but not in other bound states (pink, cyan and green). **C)** A superimposed crystal structure of L/F transferase wild-type bound with puromycin (yellow, PDB ID: 2DPT) and substrate analogue rA-Phe (green, PDB ID: 2Z3K). The space-filling dotted spheres emphasize the steric clash between M144 and puromycin.

Figure 4-5C compares the space-filling surfaces of M144 and puromycin between the puromycin-bound structure (yellow) and rA-Phe-bound structure (green). In the puromycin-bound structure (yellow), the space-filling surface of the sulphur atom of M144 is in contact with the oxygen atom of *p*-methoxy group of puromycin. On the other hand in the rA-Phe-bound structure (green), space-filling surface between the sulphur atom of M144 and the oxygen atom of *p*-methoxy group of puromycin are overlapping. This suggests that the side chain of M144 in the puromycin-bound structure (yellow) must be rotated to accommodate the *p*-methoxy group. It is also predicted that the lone pair repulsion between the sulphur atom of M144 and the oxygen atom of *p*-methoxy group further destabilizes the binding of puromycin. From this analysis of the X-ray crystal structures of L/F transferase, we propose that puromycin binding is weakened as a result of requiring M144 to adopt an alternative conformation in order to accommodate the amino acid side chain of puromycin in the binding site. This model predicts that reducing the size of the M144 side chain may selectively enhance puromycin binding.

4.2.4. Mutagenesis of M144

Based on the above observations, we hypothesized that removing steric clashes between puromycin and M144 of L/F transferase through site directed mutagenesis may increase the binding and thus inhibition of puromycin, but the same mutation is not expected to significantly alter rA-Phe-amide binding. To evaluate this hypothesis, several M144 mutations

were generated. A protein-protein basic local alignment search (blastp) was performed to determine the natural variation in sequences at position 144 of L/F transferase. Blastp searched against 100 related protein sequences revealed that methionine occurs at position 144 at a high frequency (97%) while variations such as isoleucine occur rarely (2%). Since blastp aims to search for similar protein sequences, an additional BLAST search was performed. Position-specific iterative BLAST (PSI-BLAST), which searches for distantly related proteins, was used to identify the variation at position 144 of L/F transferase from more diverse organisms. Of the 485 protein sequences searched, leucine (57.5%) and methionine (36.9%) at position 144 account for majority of the variations. Other variations include isoleucine (2.9%), phenylalanine (1.6%), valine (0.6%) and cysteine (0.2%). **Figure 4-6** summarizes the natural variation of amino acid sequences and their frequencies at position 144 of L/F transferase from other prokaryotic species as sequence logos (Schneider and Stephens 1990) (WebLogo (Crooks *et al.* 2004) Version 2.8.2).

Based on the natural variation of sequences at position 144 of L/F transferase, a series of M144 mutants were generated: M144A, M144V, M144I, M144L, and M144F. These M144 mutants vary in side chain size which makes them ideal for probing the steric effect of M144 and puromycin. The M144A mutation, which expands the deep C-shaped amino acid hydrophobic pocket in L/F transferase, had previously been reported to enable the transfer of larger unnatural amino acids to

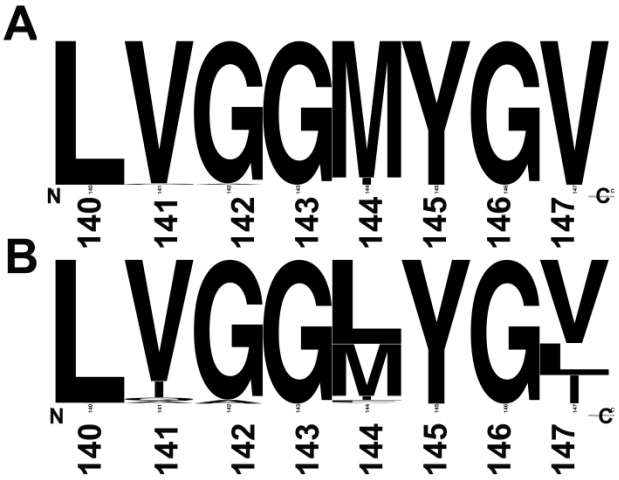

Figure 4-6: Sequence logos for the natural variation of amino acid sequences and its frequencies in position 144 of L/F transferase using A) blastp and B) PSI-BLAST.

polypeptide substrates (Taki *et al.* 2008). Thus we hypothesize that the M144A mutant would result in enhanced binding of puromycin and therefore a more significant inhibition of L/F transferase. Meanwhile the M144V, M144I, M144L, and M144F mutations were also analyzed as these are analogues of the natural sequences variation at this site for aa-transferase from other organisms.

4.2.5. Puromycin inhibition of M144 mutants

To first evaluate the effect of the M144 mutations on L/F transferase activity, enzymatic reactions of wild-type and M144 mutants were performed under standard conditions in the absence of puromycin. The determined apparent kinetic parameters are listed in **Table 4-2** and summarized in **Figure 4-7A**. The data demonstrates that M144I and M144L result in minimal loss in enzymatic activity when compared to the wild-type enzyme. Meanwhile M144A, M144V, and M144F result in less than a 2-fold loss in enzymatic activity. Together all M144 mutants result in less than a 2-fold loss in both apparent K_M and activity. This drop in activity in M144A is comparable to other active site mutants we (Fung *et al.* 2011) and others (Suto *et al.* 2006) have previously investigated.

The initial reaction rates measured for the effects of puromycin inhibition on M144 mutants are listed in **Table 4-3**, while this data is summarized in the graphs shown in **Figure 4-7**. Of all M144 mutants, the M144A mutant is most strongly inhibited by puromycin. Initial

Table 4-2: Kinetic parameters of L/F transferase catalyzed peptide bond formation by M144 mutants.

L/F transferase	Apparent K_M (peptide) (μM)	Apparent k_{cat} (min^{-1})	Catalytic Efficiency (k_{cat}^{mut}/K_M^{mut} / k_{cat}^{WT}/K_M^{WT})
Wild-type	0.47 ± 0.15	0.0192 ± 0.0008	1.0
M144A	0.51 ± 0.10	0.0126 ± 0.0005	0.604
M144V	0.71 ± 0.15	0.0097 ± 0.0004	0.334
M144I	0.77 ± 0.12	0.0156 ± 0.0005	0.496
M144L	0.54 ± 0.10	0.0155 ± 0.0005	0.703
M144F	0.46 ± 0.10	0.0105 ± 0.0004	0.559

Errors represented are the standard deviation of three independent experiments.

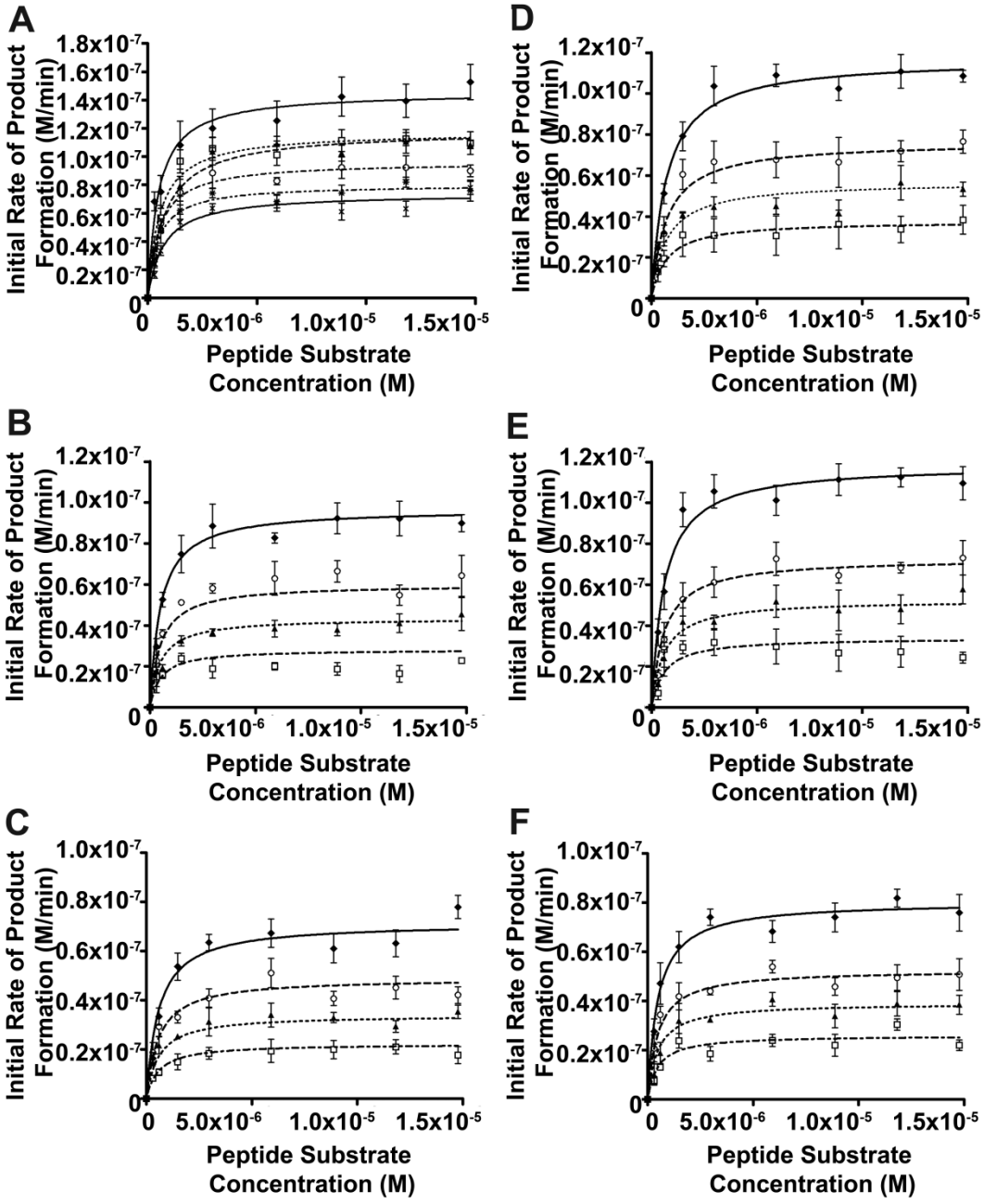
A
Initial Rate of Product Formation (M/min)
1.8x10-7
1.6x10-7
1.4x10-7
1.2x10-7
1.0x10-7
0.8x10-7
0.6x10-7
0.4x10-7
0.2x10-7
0
0 5.0x10-6 1.0x10-5 1.5x10-5
Peptide Substrate Concentration (M)
D
Initial Rate of Product Formation (M/min)
1.2x10-7
1.0x10-7
0.8x10-7
0.6x10-7
0.4x10-7
0.2x10-7
0
0 5.0x10-6 1.0x10-5 1.5x10-5
Peptide Substrate Concentration (M)
B
Initial Rate of Product Formation (M/min)
1.2x10-7
1.0x10-7
0.8x10-7
0.6x10-7
0.4x10-7
0.2x10-7
0
0 5.0x10-6 1.0x10-5 1.5x10-5
Peptide Substrate Concentration (M)
E
Initial Rate of Product Formation (M/min)
1.2x10-7
1.0x10-7
0.8x10-7
0.6x10-7
0.4x10-7
0.2x10-7
0
0 5.0x10-6 1.0x10-5 1.5x10-5
Peptide Substrate Concentration (M)
C
Initial Rate of Product Formation (M/min)
1.0x10-7
0.8x10-7
0.6x10-7
0.4x10-7
0.2x10-7
0
0 5.0x10-6 1.0x10-5 1.5x10-5
Peptide Substrate Concentration (M)
F
Initial Rate of Product Formation (M/min)
1.0x10-7
0.8x10-7
0.6x10-7
0.4x10-7
0.2x10-7
0
0 5.0x10-6 1.0x10-5 1.5x10-5
Peptide Substrate Concentration (M)

Figure 4-7: Puromycin inhibition on L/F transferase M144 mutants. **A)** A graphical comparison of enzymatic activity between wild-type L/F transferase (♦), M144A (○), M144V (x), M144I (▲), M144L (□), and M144F (*) mutants across various peptide substrate concentration and in the absence of puromycin. **B)** A graphical display of initial rate of product formation against peptide substrate concentration for L/F transferase M144A mutant in the presence of 0.0 (♦), 0.024 (○), 0.048 (▲), and 0.095 mM (□) of puromycin. Errors represented are the standard deviation of three independent experiments. **C)** A graphical display of initial rate of product formation against peptide substrate concentration for L/F transferase M144V mutant in the presence of 0.0 (♦), 0.10 (○), 0.24 (▲), and 0.48 mM (□) of puromycin. Errors represented are the standard deviation of three independent experiments. **D) - F)** A graphical display of initial rate of product formation against substrate concentration for L/F transferase **D)** M144I, **E)** M144L, and **F)** M144F mutant in the presence of 0.0 (♦), 0.24 (○), 0.48 (▲), and 0.95 mM (□) of puromycin. Errors represented are the standard deviation of three independent experiments.

Table 4-3: **Initial reaction rates of L/F transferase M144 mutants catalyzed peptide bond formation during puromycin and rA-Phe-amide inhibition.**

A) L/F transferase M144A and puromycin

L/F transferase M144A	Substrate Peptide Conc. (μM)	Initial Reaction Rate (μM min^{-1})
M144A 0.0 mM puromycin	0.30	0.030 ± 0.007
	0.59	0.053 ± 0.006
	1.5	0.075 ± 0.016
	3.0	0.089 ± 0.019
	5.9	0.083 ± 0.004
	8.9	0.092 ± 0.013
	11.8	0.092 ± 0.015
	14.8	0.090 ± 0.007
M144A 0.024 mM puromycin	0.30	0.023 ± 0.002
	0.59	0.036 ± 0.003
	1.5	0.051 ± 0.002
	3.0	0.058 ± 0.004
	5.9	0.063 ± 0.014
	8.9	0.066 ± 0.009
	11.8	0.055 ± 0.009
	14.8	0.064 ± 0.017
M144A 0.048 mM puromycin	0.30	0.018 ± 0.019
	0.59	0.019 ± 0.003
	1.5	0.032 ± 0.004
	3.0	0.037 ± 0.003
	5.9	0.038 ± 0.007
	8.9	0.038 ± 0.005
	11.8	0.041 ± 0.008
	14.8	0.046 ± 0.014
M144A 0.095 mM puromycin	0.30	0.015 ± 0.008
	0.59	0.016 ± 0.003
	1.5	0.024 ± 0.004
	3.0	0.019 ± 0.008
	5.9	0.020 ± 0.003
	8.9	0.019 ± 0.005
	11.8	0.016 ± 0.007
	14.8	0.023 ± 0.002

Errors represented are the standard deviation of three independent experiments (n=3).

B) L/F transferase M144V and puromycin

L/F transferase M144V	Substrate Peptide Conc. (μM)	Initial Reaction Rate (μM min^{-1})
M144V 0.0 mM puromycin	0.30	0.016 ± 0.004
	0.59	0.033 ± 0.006
	1.5	0.054 ± 0.010
	3.0	0.064 ± 0.006
	5.9	0.067 ± 0.010
	8.9	0.061 ± 0.010
	11.8	0.063 ± 0.010
	14.8	0.078 ± 0.008
M144V 0.10 mM puromycin	0.30	0.019 ± 0.007
	0.59	0.029 ± 0.007
	1.5	0.033 ± 0.004
	3.0	0.041 ± 0.006
	5.9	0.051 ± 0.010
	8.9	0.041 ± 0.005
	11.8	0.045 ± 0.008
	14.8	0.042 ± 0.006
M144V 0.24 mM puromycin	0.30	0.014 ± 0.001
	0.59	0.022 ± 0.006
	1.5	0.0253 ± 0.0001
	3.0	0.031 ± 0.010
	5.9	0.034 ± 0.008
	8.9	0.033 ± 0.004
	11.8	0.029 ± 0.004
	14.8	0.035 ± 0.005
M144V 0.48 mM puromycin	0.30	0.0083 ± 0.0010
	0.59	0.011 ± 0.002
	1.5	0.015 ± 0.006
	3.0	0.018 ± 0.004
	5.9	0.019 ± 0.008
	8.9	0.020 ± 0.006
	11.8	0.021 ± 0.005
	14.8	0.018 ± 0.006

Errors represented are the standard deviation of three independent experiments (n=3).

C) L/F transferase M144I and puromycin

L/F transferase M144I	Substrate Peptide Conc. (μM)	Initial Reaction Rate (μM min^{-1})
M144I 0.0 mM puromycin	0.30	0.025 ± 0.006
	0.59	0.051 ± 0.010
	1.5	0.079 ± 0.014
	3.0	0.10 ± 0.02
	5.9	0.11 ± 0.01
	8.9	0.10 ± 0.01
	11.8	0.11 ± 0.02
	14.8	0.109 ± 0.006
M144I 0.24 mM puromycin	0.30	0.020 ± 0.003
	0.59	0.032 ± 0.007
	1.5	0.061 ± 0.013
	3.0	0.067 ± 0.018
	5.9	0.068 ± 0.015
	8.9	0.066 ± 0.021
	11.8	0.072 ± 0.009
	14.8	0.077 ± 0.010
M144I 0.48 mM puromycin	0.30	0.015 ± 0.003
	0.59	0.033 ± 0.012
	1.5	0.041 ± 0.003
	3.0	0.045 ± 0.008
	5.9	0.045 ± 0.007
	8.9	0.042 ± 0.002
	11.8	0.057 ± 0.015
	14.8	0.053 ± 0.006
M144I 0.95 mM puromycin	0.30	0.013 ± 0.009
	0.59	0.020 ± 0.007
	1.5	0.031 ± 0.018
	3.0	0.031 ± 0.014
	5.9	0.031 ± 0.017
	8.9	0.036 ± 0.021
	11.8	0.034 ± 0.011
	14.8	0.038 ± 0.012

Errors represented are the standard deviation of three independent experiments (n=3).

D) L/F transferase M144L and puromycin

L/F transferase M144L	Substrate Peptide Conc. (μM)	Initial Reaction Rate (μM min^{-1})
M144L 0.0 mM puromycin	0.30	0.037 ± 0.013
	0.59	0.057 ± 0.018
	1.5	0.10 ± 0.02
	3.0	0.11 ± 0.02
	5.9	0.10 ± 0.01
	8.9	0.11 ± 0.02
	11.8	0.11 ± 0.01
	14.8	0.11 ± 0.02
M144L 0.24 mM puromycin	0.30	0.016 ± 0.004
	0.59	0.033 ± 0.002
	1.5	0.053 ± 0.014
	3.0	0.061 ± 0.013
	5.9	0.073 ± 0.014
	8.9	0.064 ± 0.007
	11.8	0.068 ± 0.005
	14.8	0.073 ± 0.015
M144L 0.48 mM puromycin	0.30	0.012 ± 0.002
	0.59	0.024 ± 0.009
	1.5	0.042 ± 0.013
	3.0	0.042 ± 0.007
	5.9	0.052 ± 0.016
	8.9	0.047 ± 0.020
	11.8	0.048 ± 0.014
	14.8	0.058 ± 0.014
M144L 0.95 mM puromycin	0.30	0.007 ± 0.005
	0.59	0.028 ± 0.023
	1.5	0.029 ± 0.005
	3.0	0.032 ± 0.011
	5.9	0.030 ± 0.015
	8.9	0.027 ± 0.016
	11.8	0.027 ± 0.013
	14.8	0.024 ± 0.005

Errors represented are the standard deviation of three independent experiments (n=3).

E) L/F transferase M144F and puromycin

L/F transferase M144F	Substrate Peptide Conc. (μM)	Initial Reaction Rate (μM min^{-1})
M144F 0.0 mM puromycin	0.30	0.028 ± 0.010
	0.59	0.047 ± 0.017
	1.5	0.062 ± 0.013
	3.0	0.074 ± 0.007
	5.9	0.068 ± 0.009
	8.9	0.074 ± 0.012
	11.8	0.082 ± 0.007
	14.8	0.076 ± 0.015
M144F 0.24 mM puromycin	0.30	0.021 ± 0.010
	0.59	0.034 ± 0.006
	1.5	0.042 ± 0.010
	3.0	0.044 ± 0.002
	5.9	0.054 ± 0.004
	8.9	0.046 ± 0.006
	11.8	0.050 ± 0.009
	14.8	0.051 ± 0.011
M144F 0.48 mM puromycin	0.30	0.010 ± 0.008
	0.59	0.019 ± 0.009
	1.5	0.032 ± 0.014
	3.0	0.032 ± 0.001
	5.9	0.041 ± 0.006
	8.9	0.034 ± 0.010
	11.8	0.039 ± 0.010
	14.8	0.039 ± 0.008
M144F 0.95 mM puromycin	0.30	0.0077 ± 0.0028
	0.59	0.013 ± 0.002
	1.5	0.024 ± 0.005
	3.0	0.019 ± 0.005
	5.9	0.024 ± 0.004
	8.9	0.022 ± 0.008
	11.8	0.030 ± 0.004
	14.8	0.022 ± 0.004

Errors represented are the standard deviation of three independent experiments (n=3).

F) L/F transferase M144A and rA-Phe-amide

L/F transferase	Substrate Peptide Conc. (μM)	Initial Reaction Rate (μM min^{-1})
M144A 0.0 mM rA-Phe-amide	0.59	0.044 ± 0.009
	3.0	0.058 ± 0.002
	5.9	0.059 ± 0.009
	11.8	0.054 ± 0.003
M144A 0.24 mM rA-Phe-amide	0.59	0.024 ± 0.003
	3.0	0.032 ± 0.004
	5.9	0.038 ± 0.006
	11.8	0.035 ± 0.014
M144A 0.48 mM rA-Phe-amide	0.59	0.017 ± 0.002
	3.0	0.024 ± 0.002
	5.9	0.026 ± 0.002
	11.8	0.023 ± 0.005
M144A 0.95 mM rA-Phe-amide	0.59	0.015 ± 0.003
	3.0	0.010 ± 0.005
	5.9	0.019 ± 0.006
	11.8	0.017 ± 0.003

Errors represented are the standard deviation of three independent experiments (n=3).

investigations of the M144A mutant using puromycin concentrations identical to that used for the wild-type enzyme revealed strong inhibition that could not be accurately quantified (data not shown). Analysis was then repeated using lower concentrations of puromycin (0, 0.024, 0.048, and 0.095 mM) on the M144A mutant (**Figure 4-7B**). The resulting apparent K_i of puromycin for L/F transferase M144A mutant is calculated to be 39 ± 3 μM (GraphPad Prism Version 5.02). When compared to 425 μM apparent K_i for wild-type L/F transferase, a 10.9-fold increase in puromycin inhibition potency is observed with the M144A mutation.

The M144V mutation (**Figure 4-7C**) also results in a slightly larger hydrophobic amino acid binding pocket. Initial investigations of the M144V mutant using puromycin concentrations identical to that used for the wild-type enzyme also revealed strong inhibition, the analysis was therefore repeated using lower concentrations of puromycin (0, 0.10, 0.24, and 0.48 mM). The M144V mutant has an apparent K_i of 218 ± 16 μM for puromycin (GraphPad Prism Version 5.02), which is a modest 2-fold increase in inhibition potency compared to the wild-type enzyme. Conversely the M144I (**Figure 4-7D**), M144L (**Figure 4-7E**), and M144F (**Figure 4-7F**) mutants, which have similar side chain size to the wild-type enzyme, have apparent inhibition constants of 452 ± 33 μM, 381 ± 28 μM and 455 ± 33 μM respectively. The apparent K_i are summarized in **Table 4-4**. The M144I, M144L, and M144F mutations do not significantly enhance puromycin inhibition. Thus, mutations introducing a larger

Table 4-4: Apparent inhibition constants (K_i) of puromycin and rA-Phe-amide for L/F transferase wild-type and M144 mutants.

L/F transferase	Apparent K_i for puromycin (μM)	Fold Change of apparent K_i for puromycin compared to wild-type	Apparent K_i for rA-Phe-amide (μM)	Fold Change of apparent K_i for rA-Phe-amide compared to wild-type
Wild-type	425 ± 36	1.0	659 ± 72	1.0
M144A	39 ± 3	10.9	352 ± 28	1.9
M144V	218 ± 16	1.9	N.D.	N.D.
M144I	452 ± 33	0.9	N.D.	N.D.
M144L	381 ± 28	1.1	N.D.	N.D.
M144F	455 ± 33	0.9	N.D.	N.D.

Errors represented are the standard deviation of three independent experiments.

N.D. - Not Determined

hydrophobic amino acid binding pocket in L/F transferase (i.e. M144A) enhanced potent inhibition by puromycin. This observation agrees with previous data that demonstrated a M144A mutation in L/F transferase enabled the transfer of large unnatural amino acids from aa-tRNAs to polypeptide substrates (Taki *et al.* 2008).

4.2.6. rA-Phe-amide inhibition on M144A

Inhibition of the M144A mutant by rA-Phe-amide was also investigated. We hypothesized that the increase in amino acid binding pocket size in L/F transferase is not expected to significantly alter rA-Phe-amide binding. A graph of initial rate of product formation against peptide substrate concentrations is shown for L/F transferase M144A (**Figure 4-8**) in the presence of 0, 0.24, 0.47, and 0.94 mM rA-Phe-amide. Initial reaction rates are summarized in **Table 4-3**. The apparent K_i for rA-Phe-amide for M144A mutant is calculated to be 352 ± 28 μM (GraphPad Prism Version 5.02). Comparing the apparent K_i for rA-Phe-amide between wild-type (659 μM) and M144A mutant (352 μM), there is a 1.9-fold increase in inhibition. Although modest, this increase in inhibition cannot be explained by the current structural and biochemical data. On the other hand, the increased inhibition achieved by L/F transferase M144A mutant with rA-Phe-amide (1.9-fold) was significantly smaller than the increased inhibition achieved by the M144A mutant with puromycin (10.9-fold). This suggests that, as expected, steric hindrance plays a larger effect with puromycin than with rA-Phe-amide.

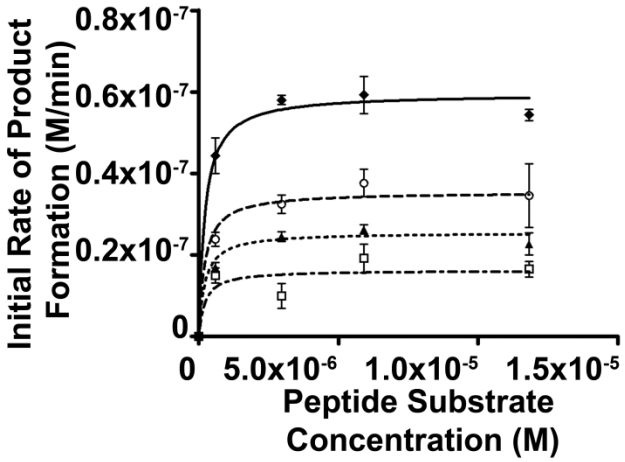

Figure 4-8: rA-Phe-amide Inhibition on L/F transferase M144A mutant. A graphical display of initial rate of product formation versus peptide substrate concentration for L/F transferase M144A mutant enzyme in the presence of 0.0 (♦), 0.24 (○), 0.47 (▲), and 0.94 mM (□) rA-Phe-amide. The lack of contacts may be the cause of poor binding by rA-Phe-amide on the M144A mutant. Errors represented are the standard deviation of three independent experiments.

4.3. Discussion

4.3.1. A proposed model for puromycin and rA-Phe-amide mechanism of binding and inhibition

Table 4-4 shows a summary of all the apparent K_i determined. A comparison of the apparent K_i of puromycin (425 μM) and rA-Phe-amide (659 μM) for wild-type L/F transferase shows that both substrate analogues bind to L/F transferase with similar order of affinities. However the substrate analogue-bound X-Ray crystal structures indicated that both analogues bind differently. We hypothesized that the presence of a modified adenine base has opposite effects on binding when compared to the modified phenylalanine, which serendipitously results in similar binding affinity.

To investigate the role of *p*-methoxyphenylalanine group of puromycin on binding, we demonstrated that reducing the size of the M144 side chain selectively enhances puromycin binding to L/F transferase. Recall that the side chain of M144 of L/F transferase is rotated and pointed away from the *p*-methoxyphenylalanine group of puromycin to avoid steric clashes (**Figure 4-5B**). Mutations which introduce a larger hydrophobic amino acid binding pocket in L/F transferase, such as M144A, enabled stronger inhibition effects by puromycin. Mutations that retained the hydrophobic pocket size, such as M144I, M144L, and M144F retained the poor inhibition effects by puromycin. Additionally, as expected, rA-Phe-amide inhibition on the

M144A mutant enzyme is not as significantly improved as with puromycin perhaps due to a lack of enzyme-analogue contacts.

4.3.2. Concluding Remarks

Here we were able to rationalize and demonstrate that the differences in both the substrate analogue-bound X-ray crystal structures are real and that the similar binding affinities between the analogues is somewhat serendipitous as they both use a series of different interactions. Through structural analysis, mutagenesis, and enzymatic activity assays, we have determined that the poor binding of puromycin can be greatly increased by enlarging the size of the hydrophobic binding pocket for the amino acid side chain. Thus, the modified adenine base and modified phenylalanine play opposite roles in binding. The modified adenine base has favourable interactions with L/F transferase via π-π stacking and additional hydrophobic interactions, whereas the modified phenylalanine hinders binding via a steric clash with M144 in the C-shaped hydrophobic pocket. Based on the results, we hypothesize that a substrate analogue of rA-Phe-amide with modified base may inhibit and bind to L/F transferase more strongly. Alternatively, significantly weaker binding of puromycin and rA-Phe-amide may be a result of a lack of extensive contacts between the small molecule analogue and L/F transferase, which is probably present for the natural aa-tRNA substrate.

4.4. References

Abramochkin, G., and Shrader, T.E. (1996) Aminoacyl-tRNA recognition by the leucyl/phenylalanyl-tRNA-protein transferase. *J.Biol.Chem.* **271,** 22901-22907

Baker, R.T., and Varshavsky, A. (1991) Inhibition of the N-end rule pathway in living cells. *Proc.Natl.Acad.Sci.U.S.A.* **88,** 1090-1094

Berleth, E.S., Li, J., Braunscheidel, J.A., and Pickart, C.M. (1992) A reactive nucleophile proximal to vicinal thiols is an evolutionarily conserved feature in the mechanism of Arg aminoacyl-tRNA protein transferase. *Arch.Biochem.Biophys.* **298,** 498-504

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004) WebLogo: a sequence logo generator. *Genome Res.* **14,** 1188-1190

Dong, X., Kato-Murayama, M., Muramatsu, T., Mori, H., Shirouzu, M., Bessho, Y., and Yokoyama, S. (2007) The crystal structure of leucyl/phenylalanyl-tRNA-protein transferase from *Escherichia coli. Protein Sci.* **16,** 528-534

Ebhardt, H.A., Xu, Z., Fung, A.W., and Fahlman, R.P. (2009) Quantification of the post-translational addition of amino acids to proteins by MALDI-TOF mass spectrometry. *Anal.Chem.* **81,** 1937-1943

Ennis, H.L. (1965a) Inhibition of protein synthesis by polypeptide antibiotics. I. Inhibition in intact bacteria. *J.Bacteriol.* **90,** 1102-1108

Ennis, H.L. (1965b) Inhibition of protein synthesis by polypeptide antibiotics. II. *in vitro* protein synthesis. *J.Bacteriol.* **90,** 1109-1119

Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature.* **347,** 203-206

Fung, A.W., Ebhardt, H.A., Abeysundara, H., Moore, J., Xu, Z., and Fahlman, R.P. (2011) An alternative mechanism for the catalysis of peptide bond formation by L/F transferase: substrate binding and orientation. *J.Mol.Biol.* **409,** 617-629

Horinishi, H., Hashizume, S., Seguchi, M., and Takahashi, K. (1975) Incorporation of methionine by a soluble enzyme system from *Escherichia coli. Biochem.Biophys.Res.Commun.* **67,** 1136-1143

Hu, R.G., Wang, H., Xia, Z., and Varshavsky, A. (2008) The N-end rule pathway is a sensor of heme. *Proc.Natl.Acad.Sci.U.S.A.* **105,** 76-81

Klemperer, N.S., and Pickart, C.M. (1989) Arsenite inhibits two steps in the ubiquitin-dependent proteolytic pathway. *J.Biol.Chem.* **264,** 19245-19252

Laidler, K. (1955) Theory of the Transient Phase in Kinetics, with Special Reference to Enzyme Systems. *Canadian Journal of Chemistry-Revue Canadienne De Chimie.* **33,** 1614-1624

Leibowitz, M.J., and Soffer, R.L. (1970) Enzymatic modification of proteins. III. Purification and properties of a leucyl, phenylalanyl transfer ribonucleic acid protein transferase from *Escherichia coli. J.Biol.Chem.* **245,** 2066-2073

Leibowitz, M.J., and Soffer, R.L. (1971) Enzymatic modification of proteins. VII. Substrate specificity of leucyl,phenylalanyl-transfer ribonucleic acid-protein transferase. *J.Biol.Chem.* **246,** 5207-5212

Li, J., and Pickart, C.M. (1995a) Binding of phenylarsenoxide to Arg-tRNA protein transferase is independent of vicinal thiols. *Biochemistry.* **34,** 15829-15837

Li, J., and Pickart, C.M. (1995b) Inactivation of arginyl-tRNA protein transferase by a bifunctional arsenoxide: identification of residues proximal to the arsenoxide site. *Biochemistry.* **34,** 139-147

Palmer, T. (1991) Understanding Enzymes Third Edition. Ellis Horwood Limited.

Schneider, T.D., and Stephens, R.M. (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* **18,** 6097-6100

Schnell, S., and Maini, P.K. (2000) Enzyme kinetics at high enzyme concentration. *Bull.Math.Biol.* **62,** 483-499

Starck, S.R., and Roberts, R.W. (2002) Puromycin oligonucleotides reveal steric restrictions for ribosome entry and multiple modes of translation inhibition. *RNA.* **8,** 890-903

Suto, K., Shimizu, Y., Watanabe, K., Ueda, T., Fukai, S., Nureki, O., and Tomita, K. (2006) Crystal structures of leucyl/phenylalanyl-tRNA-protein transferase and its complex with an aminoacyl-tRNA analog. *EMBO J.* **25,** 5942-5950

Taki, M., Kuroiwa, H., and Sisido, M. (2008) Chemoenzymatic transfer of fluorescent non-natural amino acids to the N terminus of a protein/peptide. *Chembiochem.* **9,** 719-722

Wagner, A.M., Fegley, M.W., Warner, J.B., Grindley, C.L., Marotta, N.P., and Petersson, E.J. (2011) N-terminal protein modification using simple aminoacyl transferase substrates. *J.Am.Chem.Soc.* **133,** 15139-15147

Watanabe, K., Toh, Y., Suto, K., Shimizu, Y., Oka, N., Wada, T., and Tomita, K. (2007) Protein-based peptide-bond formation by aminoacyl-tRNA protein transferase. *Nature.* **449,** 867-871

Yang, F., Xia, X., Lei, H.Y., and Wang, E.D. (2010) Hemin binds to human cytoplasmic arginyl-tRNA synthetase and inhibits its catalytic activity. *J.Biol.Chem.* **285,** 39437-39446

Chapter 5

The Characterization of Aminoacyl-tRNA Recognition by L/F transferase

A version of this chapter is published in:
Fung AW, Leung CC and Fahlman RP (2014) The Determination of $tRNA^{Leu}$ Recognition Nucleotides for *Escherichia coli* L/F transferase. *RNA*. 20 (8): 1210-1222.

5.1. Introduction

Transfer RNAs (tRNAs), in addition to their prominent role in translation, also participate in alternative functions in many organisms *in vivo* including amino acid biosynthesis (Ibba and Soll 2004, Sheppard *et al.* 2008), antibiotic biosynthesis (Nolan and Walsh 2009), cell envelope remodeling (Villet *et al.* 2007, Roy and Ibba 2008, Fonvielle *et al.* 2009, Giannouli *et al.* 2009), and targeted proteolysis (Abramochkin and Shrader 1996, Mogk *et al.* 2007) (see review for more details (Banerjee *et al.* 2010, Francklyn and Minajigi 2010)). These alternative functions often utilize aa-tRNA as a source of activated amino acids, yet deacyl-tRNAs also have regulatory roles in gene expression during amino acid starvation in both prokaryotes and eukaryotes (Wendrich *et al.* 2002, Zaborske *et al.* 2009).

Currently, there is a conundrum regarding the precise *in vivo* mechanism in which the aa-tRNA used for alternative functions evades the protein biosynthesis machinery. Elongation factor Tu (EF-Tu) binds to an aa-tRNA molecule in the cytoplasm where it hydrolyzes a GTP molecule and releases the aa-tRNA to the ribosomal A-site for protein synthesis (Marshall *et al.* 2008, Agirrezabala and Frank 2009, Schmeing

and Ramakrishnan 2009). The *in vivo* concentration of total aa-tRNA and EF-Tu are comparable (~100 μM) and EF-Tu binds to all aa-tRNAs with strong, similar affinities (K_D in low nM range) (Andersen and Wiborg 1994, LaRiviere *et al.* 2001, Schrader *et al.* 2011). Some possible evasion mechanisms have been described, such as channeling substrates through a protein complex with aminoacyl-tRNA synthetases (Bailly *et al.* 2007), regulating the subcellular localization of tRNAs (Stortchevoi *et al.* 2003), having competitive binding affinities to aa-tRNAs (Roy and Ibba 2008), utilizing misacylated tRNAs (Stanzel *et al.* 1994, Becker and Kern 1998) or idiosyncratic features of specific tRNA isoacceptors (Giannouli *et al.* 2009). However, it remains unclear how free, canonical aa-tRNA species participate in both translation and alternative functions.

L/F transferase catalyzes the post-translational addition of amino acids using aa-tRNA as donor substrates, which targets the modified acceptor protein substrate for degradation via the N-end rule (Leibowitz and Soffer 1969, Tobias *et al.* 1991). L/F transferase has degenerate aa-tRNA specificity *in vitro* where it utilizes Leu-tRNA^{Leu}, Phe-tRNA^{Phe} (Leibowitz and Soffer 1969), and to a lesser extent Met-tRNA^{Met} as

substrates (Scarpulla *et al.* 1976, Abramochkin and Shrader 1996). *In vivo* studies however suggest that leucylation is the dominant modification (Shrader *et al.* 1993).

In vitro studies with misacylated tRNAs (Leibowitz and Soffer 1971, Abramochkin and Shrader 1996) and minimalistic adenosine esters of natural and unnatural amino acids (3' rA-aa) (Wagner *et al.* 2011) are sufficient for aminoacyl transfer. This suggests that the major determinant of tRNA recognition by L/F transferase is the 3' terminal adenosine and the aminoacyl moiety of an aa-tRNA (Leibowitz and Soffer 1971, Abramochkin and Shrader 1996). Since there is no crystal structures solved for L/F transferase in complex with an intact aa-tRNA bound, the molecular insights derived from these structures remain within the 3' rA-aa of an aa-tRNA. The current L/F transferase tRNA recognition model includes the recognition of the 3' aminoacyl adenosine, the sequence-independent docking of the tRNA D-stem to the positively charged cluster (R76, R80, K83, R84) of L/F transferase, and the disruption or bending of the 3' acceptor stem of the tRNA during catalysis (Leibowitz and Soffer

1971, Abramochkin and Shrader 1996, Suto *et al.* 2006, Watanabe *et al.* 2007).

It has also been demonstrated that there is a preference for the $tRNA^{Leu}$ (anticodon 5'-CAG-3') isoacceptor for L/F transferase activity (Rao and Kaji 1974). The current tRNA recognition model does not explain this isoacceptor preference. Additionally, there is no data indicating the presence of a specialized LeuRS or $tRNA^{Leu}$ isoacceptor for L/F transferase. A comparison with the reported apparent K_M values for the minimal substrate phenylalanyl adenosine (rA-Phe, 124 μM) and Phe-$tRNA^{Phe}$ (~2 μM), suggests that the tRNA body contributes to L/F transferase recognition significantly (Rao and Kaji 1974, Abramochkin and Shrader 1996, Wagner *et al.* 2011). *In vitro* assays with mutant tRNAs suggest that the anticodon and variable loop are not the basis for specificity (Abramochkin and Shrader 1996). Upon meta-analysis of various tRNA recognition studies, Abramochkin *et al*. observed a correlation between the “strength” of the acceptor stem base pairs and overall L/F transferase activity, and hypothesized that tRNAs with weak acceptor stems (i.e. mismatches or more A:U base pairs) are better L/F

transferase substrates (Rao and Kaji 1974, Scarpulla *et al.* 1976, Abramochkin and Shrader 1996).

Here we present the molecular basis of the $tRNA^{Leu}$ isoacceptor specificity as a result of previously unidentified sequence elements in the acceptor stem that are recognized by L/F transferase. Using *in vitro* transcribed tRNAs and quantitative matrix assisted laser desorption/ ionization time-of-flight mass spectrometry (MALDI-ToF MS) enzyme activity assay developed by our lab (Ebhardt *et al.* 2009, Fung *et al.* 2011, Fung *et al.* 2014), our results demonstrated that there is a preference for the CAG isoacceptor by L/F transferase. Through mutations at the acceptor, D-, and T-stem of $tRNA^{Leu}$ isoacceptors, we identified two independent, sequence elements in the acceptor stem of Leu-$tRNA^{Leu}$ that are important for optimal L/F transferase binding and catalysis. These include the G_3:C_{70} base pair and a set of four nucleotides in the acceptor stem contribute to optimal tRNA recognition by L/F transferase. This study demonstrates that tRNA recognition by L/F transferase is more specific and sequence dependent than previously hypothesized.

5.2. Results

5.2.1. L/F transferase Activity Assays with $tRNA^{Leu}$, $tRNA^{Phe}$, and $tRNA^{Met}$ Isoacceptors

As the initial investigations with the different $tRNA^{Leu}$ isoacceptors was incomplete (Rao and Kaji 1974, Abramochkin and Shrader 1996), we initiated our investigations by examining all five *E. coli* $tRNA^{Leu}$ isoacceptors as L/F transferase substrates using the quantitative MALDI-ToF MS-based activity assay we previously developed (Ebhardt *et al.* 2009, Fung *et al.* 2011). Minimal differences using either purified (fully modified) or *in vitro* transcribed (unmodified) tRNAs was previously reported (Abramochkin and Shrader 1995) and with our need to generate hybrid tRNAs, we utilized *in vitro* transcribed tRNAs as substrates. Sequences reported for *E. coli* tRNA genes in the genomic tRNA data base (Chan and Lowe 2009) were used for template design. $tRNA^{Leu}$ isoacceptors (including $tRNA^{Leu}$ (anticodon 5'-CAG-3'), $tRNA^{Leu}$ (UAG), $tRNA^{Leu}$ (CAA), $tRNA^{Leu}$ (UAA), and $tRNA^{Leu}$ (GAG)) were *in vitro* transcribed and purified. In addition to the $tRNA^{Leu}$s, we also transcribed

and purified the sole $tRNA^{Phe}$ (GAA) isoacceptor and two representative elongator $tRNA^{Met}$ (CAU) sequences, which were chosen from the *ileX* and *metT* genes. We omitted the initiator $tRNA^{Met}$ as the characteristic 5'-cytosine overhang is not compatible with the T7 promoter site for *in vitro* transcription. Also, Scarpulla *et al*. suggested that L/F transferase does not utilize initiator $tRNA^{Met}$s (Scarpulla *et al.* 1976).

Figure 5-1 shows the cloverleaf structures of the tRNAs investigated. In addition to differences in the esterified amino acids of the tRNA substrates, cloverleaf structural examination reveals substantial differences. First, all $tRNA^{Leu}$ isoacceptors have larger D-loops and diverse, lengthy variable loops compared to $tRNA^{Phe}$ and $tRNA^{Met}$. Additionally, some individual differences are observed such as the A:C mismatch at the base of the acceptor stem for $tRNA^{Leu}$ (CAG). It had been previously hypothesized that L/F transferase specifically favours a Leu-$tRNA^{Leu}$ with weak base pairs in the acceptor stem (mismatches or more A:U pairs) (Abramochkin and Shrader 1996).

Enzymatic analysis of the addition of an amino acid to a peptide substrate in the presence of different tRNA isoacceptors was a modified

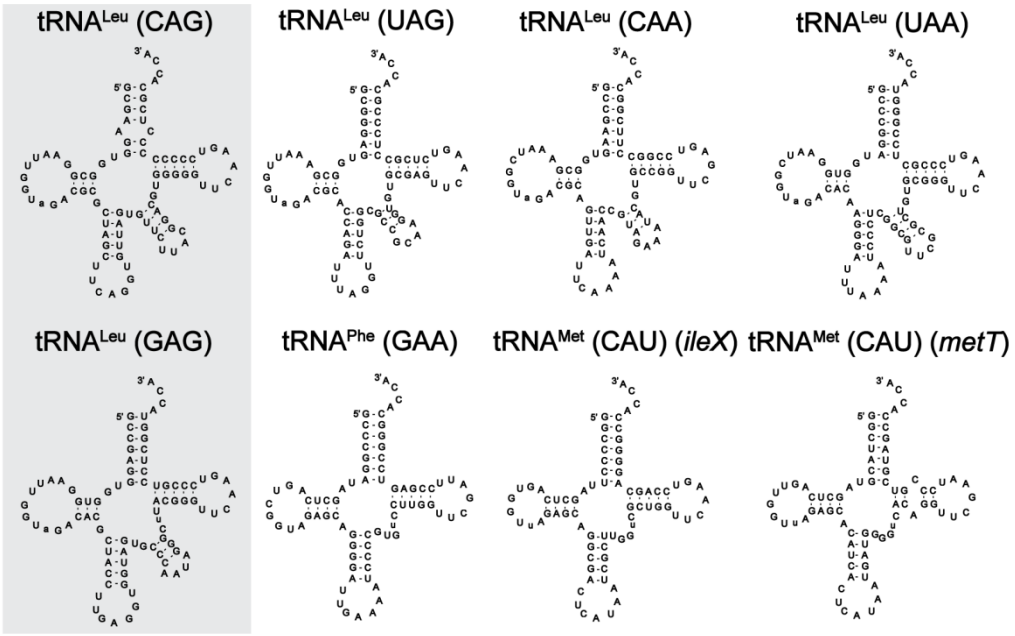

Figure 5-1. Cloverleaf structures of *E. coli* tRNA isoacceptors for leucine, phenylalanine, and methionine. The two tRNA^{Leu} isoacceptor species of importance in this study are highlighted in the grey box. Two representative elongator methionyl-tRNA species are selected for this study with their respective gene names.

method from our previously established quantitative MALDI-ToF MS method (Ebhardt *et al.* 2009, Fung *et al.* 2011, Fung *et al.* 2014), where we varied the concentration of tRNA substrate and measured product peptide formation in a continuous aminoacylation system. **Figure 5-2A** shows a graph of quantified peptide product formation over time for three representative tRNAs ($tRNA^{Leu}$ (CAG), $tRNA^{Leu}$ (GAG), and $tRNA^{Phe}$ (GAA)). Initial rates of product formation are calculated from the slope of the linear tangent line to the curve. Initial rates of product formation determined for the eight tRNA isoacceptors ($tRNA^{Leu}$, $tRNA^{Phe}$, and $tRNA^{Met}$) are listed in **Table 5-1**, while the data is summarized graphically in **Figure 5-2B** and the kinetic parameters are listed in **Table 5-2**.

Our data is in agreement with previous investigations reporting that leucylation is the optimal amino acid addition by *E. coli* L/F transferase in comparison to phenylalanylation and methionylation (Shrader *et al.* 1993). Our data in **Figure 5-2B** and **Table 5-2** demonstrates that L/F transferase is most specific to Leu-$tRNA^{Leu}$ (CAG) isoacceptor with an apparent K_M of 2.0 ± 0.4 μM with an apparent k_{cat} of 0.100 ± 0.003 min^{-1}. This is in agreement with the observed isoacceptor preference by Rao and Kaji

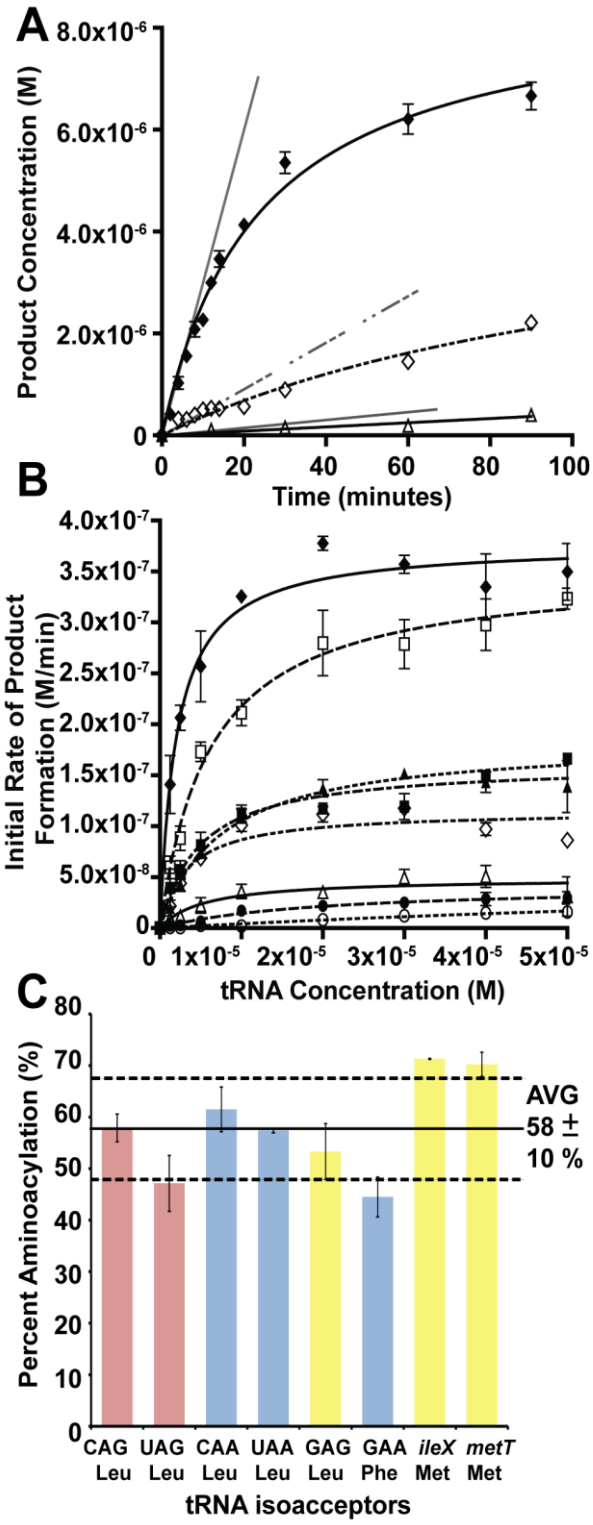
A
8.0x10-6
6.0x10-6
4.0x10-6
2.0x10-6
0
Product Concentration (M)
0
20
40
60
80
100
Time (minutes)
B
4.0x10-7
3.5x10-7
3.0x10-7
2.5x10-7
2.0x10-7
1.5x10-7
1.0x10-7
5.0x10-8
0
Initial Rate of Product Formation (M/min)
0
1x10-5
2x10-5
3x10-5
4x10-5
5x10-5
tRNA Concentration (M)
C
80
70
60
50
40
30
20
10
0
Percent Aminoacylation (%)
AVG
58 ±
10 %
CAG Leu
UAG Leu
CAA Leu
UAA Leu
GAG Leu
GAA Phe
ileX Met
metT Met
tRNA isoacceptors

Figure 5-2: A preference of leucyl-tRNA (CAG) isoacceptor by L/F transferase. **A)** Graphical analysis of product formation over time for $tRNA^{Leu}$ (CAG) (♦), $tRNA^{Phe}$ (GAA) (◊), and $tRNA^{Leu}$ (GAG) (Δ) when using an initial tRNA substrate concentration of 1.25 μM. Errors represented are standard deviation of triplicate measurements of a single independent experiment. Initial rate of product formation is calculated from the slope of the linear tangent line (grey) drawn to the curve. **B)** A graphical display of initial rate of product formation versus tRNA concentration for isoacceptors $tRNA^{Leu}$ (CAG) (♦), $tRNA^{Leu}$ (UAG) (□), $tRNA^{Leu}$ (CAA) (■), $tRNA^{Leu}$ (UAA) (▲), $tRNA^{Phe}$ (GAA) (◊), $tRNA^{Leu}$ (GAG) (Δ), $tRNA^{Met}$ (CAU) *ileX* (●), and $tRNA^{Met}$ (CAU) *metT* (○). Errors represented are the standard deviation of three independent experiments. **C)** A bar graph presenting the maximal percent aminoacylation for natural isoacceptors after 7 minutes of aminoacylation. Errors represented are the standard deviation of three independent experiments.

Table 5-1: Initial reaction rates of L/F transferase catalyzed peptide bond formation by *in vitro* transcribed *E. coli* leucyl-, phenylalanyl-, and methionyl-tRNA isoacceptors.

tRNA	tRNA Concentration (μM)	Initial Reaction Rate ($\mu M\ min^{-1}$)
$tRNA^{Leu}$ (5'-CAG-3')	1.3	0.141 ± 0.049
	2.5	0.206 ± 0.021
	5.0	0.257 ± 0.060
	10.0	0.325 ± 0.002
	20.0	0.378 ± 0.012
	30.0	0.357 ± 0.015
	40.0	0.335 ± 0.056
	50.0	0.350 ± 0.048
$tRNA^{Leu}$ (UAG)	1.3	0.060 ± 0.019
	2.5	0.088 ± 0.021
	5.0	0.173 ± 0.017
	10.0	0.211 ± 0.022
	20.0	0.280 ± 0.055
	30.0	0.279 ± 0.042
	40.0	0.298 ± 0.042
	50.0	0.323 ± 0.018
$tRNA^{Leu}$ (CAA)	1.3	0.039 ± 0.032
	2.5	0.057 ± 0.016
	5.0	0.081 ± 0.022
	10.0	0.112 ± 0.014
	20.0	0.117 ± 0.008
	30.0	0.119 ± 0.022
	40.0	0.149 ± 0.008
	50.0	0.166 ± 0.009
$tRNA^{Leu}$ (UAA)	1.3	0.020 ± 0.013
	2.5	0.043 ± 0.012
	5.0	0.074 ± 0.007
	10.0	0.110 ± 0.012
	20.0	0.137 ± 0.015
	30.0	0.152 ± 0.010
	40.0	0.143 ± 0.018
	50.0	0.139 ± 0.044
$tRNA^{Leu}$ (GAG)	1.3	0.010 ± 0.008
	2.5	0.012 ± 0.005
	5.0	0.022 ± 0.013
	10.0	0.036 ± 0.012
	20.0	0.035 ± 0.001
	30.0	0.050 ± 0.014
	40.0	0.051 ± 0.019
	50.0	0.031 ± 0.034

$tRNA^{Phe}$ (GAA)	1.3	0.022 ± 0.003
	2.5	0.044 ± 0.013
	5.0	0.069 ± 0.009
	10.0	0.101 ± 0.011
	20.0	0.112 ± 0.013
	30.0	0.118 ± 0.005
	40.0	0.097 ± 0.011
	50.0	0.086 ± 0.010
$tRNA^{Met}$ (CAU) *ileX*	1.3	0.002 ± 0.002
	2.5	0.003 ± 0.002
	5.0	0.006 ± 0.004
	10.0	0.017 ± 0.007
	20.0	0.022 ± 0.009
	30.0	0.025 ± 0.010
	40.0	0.028 ± 0.011
	50.0	0.030 ± 0.011
$tRNA^{Leu}$ (CAU) *metT*	1.3	0.001 ± 0.001
	2.5	0.001 ± 0.001
	5.0	0.003 ± 0.002
	10.0	0.002 ± 0.001
	20.0	0.008 ± 0.003
	30.0	0.012 ± 0.005
	40.0	0.015 ± 0.006
	50.0	0.016 ± 0.005

Errors represented are the standard deviation of three independent experiments.

Table 5-2: Kinetic parameters of L/F transferase catalyzed peptide bond formation by *in vitro* transcribed *E. coli* leucyl-, phenylalanyl-, and methionyl-tRNA isoacceptors.

gene	tRNA (5'-anticodon-3')	Codon (5'-3')	Apparent K_M (μM)	Apparent k_{cat} (min^{-1})	Catalytic Efficiency (k_{cat}/K_M / k_{cat}^{CAG}/K_M^{CAG})
leuPQTV	$tRNA^{Leu}$ (CAG)	CUG	2.0 ± 0.4	0.100 ± 0.003	1.0
leuW	$tRNA^{Leu}$ (UAG)	CUA, CUG	6.2 ± 1.0	0.094 ± 0.004	0.303
leuX	$tRNA^{Leu}$ (CAA)	UUG	4.9 ± 1.0	0.043 ± 0.002	0.176
leuZ	$tRNA^{Leu}$ (UAA)	UUA, UUG	7.5 ± 1.0	0.049 ± 0.002	0.131
leuU	$tRNA^{Leu}$ (GAG)	CUC, CUU	5.4 ± 3.2	0.013 ± 0.002	0.048
pheUV	$tRNA^{Phe}$ (GAA)	UUC, UUU	3.3 ± 0.7	0.031 ± 0.002	0.188
ileX	$tRNA^{Met}$ (CAU) [*ileX*]	AUG	19.3 ± 9.8	0.011 ± 0.002	0.011
metT	$tRNA^{Met}$ (CAU) [*metT*]	AUG	148 ± 188	0.018 ± 0.018	0.002

Errors represented are the standard deviation of three independent experiments.

(Rao and Kaji 1974). Although an apparent K_M of 0.11 μM has been reported for the CAG isoacceptor under different reaction conditions (Rao and Kaji 1974, Abramochkin and Shrader 1996), the relative K_M fold changes among the isoacceptors tested (CAG, UAG, CAA, and UAA) are comparable between the two studies. Subsequent data will be compared to the CAG isoacceptor tRNA as the optimal reference.

Both Met-tRNA^{Met} isoacceptors are poor substrates of L/F transferase with large apparent K_M values (9.5 and 74-fold increase) and low relative activity (90 and 500-fold decrease in catalytic efficiency). Phe-tRNA^{Phe} is an intermediate substrate with an apparent K_M of 3.3 ± 0.7 μM with mid-range relative activity (5.3-fold decrease in catalytic efficiency). Interestingly, there is a significant difference between the initial rates among the five leucine isoacceptors (**Figure 5-2B**). Specifically Leu-tRNA^{Leu} (GAG) isoacceptor, which has not been previously tested, is as poor of a substrate as tRNA^{Met} with a 21-fold decrease in catalytic efficiency when compared to the CAG isoacceptor. Since the leucine isoacceptors have equal amino acid identity contribution, comparison amongst these tRNAs directly reflects the structural contribution to the

recognition and catalytic efficiency by L/F transferase. Among the leucine isoacceptors, the apparent K_M values are within 4-fold change and the apparent k_{cat} values are within 7.6-fold change when compared to the CAG anticodon containing isoacceptor. This suggests that L/F transferase recognizes all leucine isoacceptors with similar affinities, but there is preferred recognition of certain isoacceptors over others for their catalytic efficiencies.

The catalytic efficiency differences between the leucine isoacceptors cannot be explained by the current recognition model. The current recognition model of tRNA by L/F transferase suggests that L/F transferase recognizes mainly the 3' terminal adenosine and the aminoacyl moiety, while the remaining of the tRNA body enhances binding affinity in a sequence independent manner. We hypothesize that there is a more specific recognition mechanism for tRNA binding. There are two potential explanations to the differences between CAG and GAG isoacceptors. Since our assay uses *in vitro* transcribed tRNAs, there is a possibility that some tRNAs are not folded properly and hence may not be aminoacylated properly. Alternatively, the current model regarding the

recognition requires modification, and L/F transferase recognizes beyond the 3' terminal aminoacylated adenosine.

5.2.2. Differences in Aminoacylation

To ensure that the differences in L/F transferase product formation rates is due to differences in the RNA sequence and structure but not due to reduced aminoacylation, experiments to test aminoacylation for all tRNAs were performed similarly as previously described (Wolfson and Uhlenbeck 2002). **Figure 5-2C** shows a bar graph plotting the maximal percent aminoacylation after 7 minutes for each of the tRNA isoacceptors (for full time course see **Figure 5-3**). We found that all *in vitro* transcribed tRNA isoacceptors were aminoacylated between 45-71%. Specifically, $tRNA^{Met}$s were aminoacylated to 71% and $tRNA^{Phe}$ were aminoacylated to 45%. $tRNA^{Leu}$ isoacceptors were aminoacylated to between 47-62%. Although there are variations in aminoacylation between tRNA isoaccepting species, these differences do not extrapolate to the kinetic differences observed in **Figure 5-2B** (see **Figure 5-4**). Thus, the specific

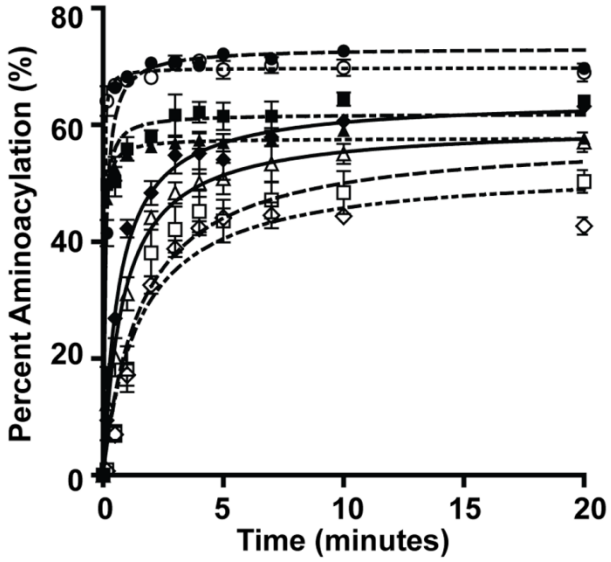

Figure 5-3: Aminoacylation time courses show that the $tRNA^{Leu}$, $tRNA^{Phe}$, and $tRNA^{Met}$ isoacceptors are efficiently aminoacylated by LeuRS, PheRS, and MetRS respectively. Percent aminoacylation versus time graph for $tRNA^{Leu}$ (CAG) (♦), $tRNA^{Leu}$ (UAG) (□), $tRNA^{Leu}$ (CAA) (■), $tRNA^{Leu}$ (UAA) (▲), $tRNA^{Phe}$ (GAA) (◊), $tRNA^{Leu}$ (GAG) (Δ), $tRNA^{Met}$ (CAU) *ileX* (●), and $tRNA^{Met}$ (CAU) *metT* (○). Errors represented are the standard deviation of three independent experiments.

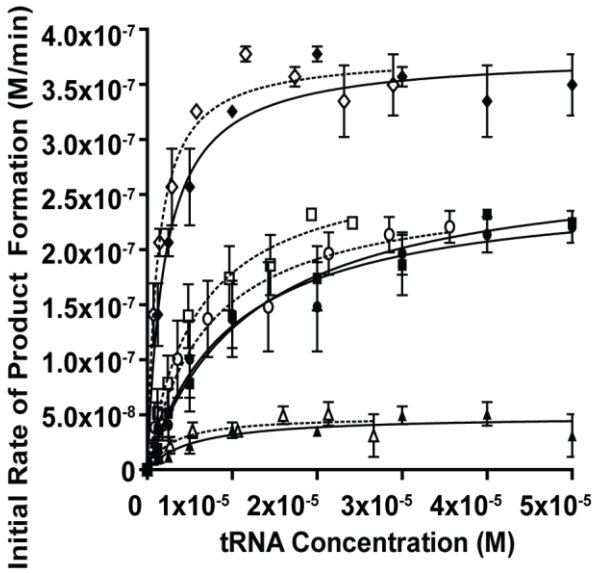

Figure 5-4: Comparison of original (filled) and corrected (open) tRNA concentration via aminoacylation assay between $tRNA^{Leu}$ (CAG) (♦, aminoacylates to 57.9% at 7 min), $tRNA^{Leu}$ (GAG) (∆, aminoacylates to 53.3%), construct 7 (♀, aminoacylates to 71.2%) and construct 8 (×, aminoacylates to 48.3%). Comparing the re-calculated values for the most varied aminoacylated constructs 7 and 8 (22.9% difference in aminoacylation), the calculated K_M values calculated alter by less than 2-fold while the k_{cat} values alter less than 1-fold. The relative k_{cat}/K_M between different isoacceptors fold changes remains similar between the original or adjusted tRNA concentrations. Errors represented are the standard deviation of three independent experiments.

substrate specificity is not due to differential aminoacylation and it is inherent to the sequence and structure of the aa-tRNA.

Our percent aminoacylation values are within the typical range observed for this method, and they reflect the equilibrium state of aminoacylation rate by aaRS and spontaneous deacylation rate (Wolfson and Uhlenbeck 2002). Since the uncharged tRNA fraction remains relatively high, we also examined whether the presence of uncharged tRNA (also a product of the reaction) significantly inhibit L/F transferase. Preliminary competition assays of uncharged $tRNA^{Phe}$ in a leucylation L/F transferase assay and uncharged $tRNA^{Leu}$ in a phenylalanylation L/F transferase assay suggest that the uncharged tRNA can compete for binding under high concentrations (**Figure 5-5**). Comparing the apparent K_i (uncharged $tRNA^{Leu}$ = 31 μM and uncharged $tRNA^{Phe}$ = 25 μM) with the apparent K_M (Leu-$tRNA^{Leu}$ = 2 μM and Phe-$tRNA^{Phe}$ = 3 μM in this study and rA-Phe = 124 μM (Wagner *et al.* 2011)) confirms that the recognition of aa-tRNA substrate by L/F transferase is through both the amino acid moiety and the tRNA body. The amino acid moiety contributes to approximately 8-fold difference (comparing Phe-$tRNA^{Phe}$ with uncharged

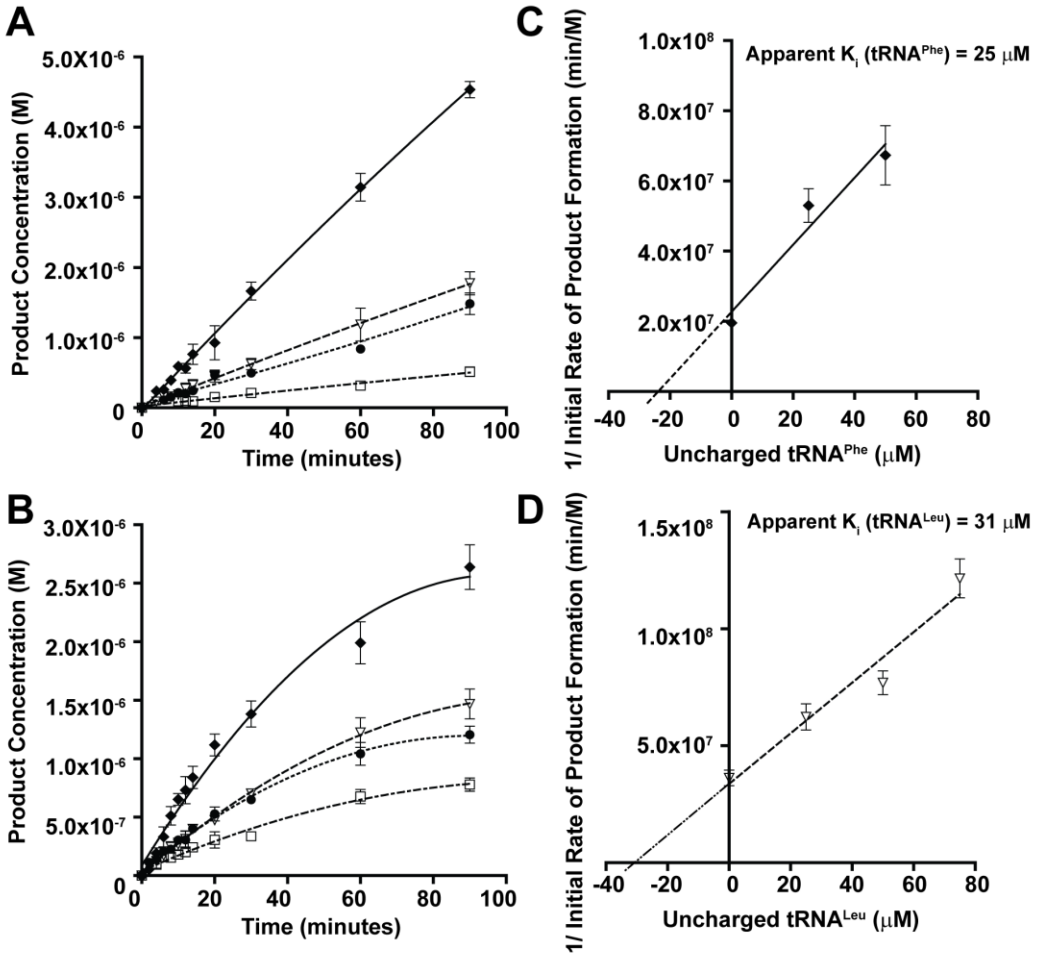

Figure 5-5: Uncharged-tRNA does not bind to L/F transferase as efficiently as aminoacylated-tRNA. **A)** Product concentration versus time graph of a leucylation L/F transferase assay with 5 μM $tRNA^{Leu}$ (CAG) and competing with 0 (♦), 25 (▽), 50 (●), and 75 (□) μM of uncharged $tRNA^{Phe}$. **B)** Product concentration versus time graph of a phenylalanylation L/F transferase assay with 5 μM $tRNA^{Phe}$ (GAA) and competing with 0 (♦), 25 (▽), 50 (●), and 75 (□) μM of uncharged $tRNA^{Leu}$. To determine the apparent K_i of uncharged tRNA to L/F transferase, tertiary plots for **C)** uncharged $tRNA^{Phe}$ (♦) and **D)** uncharged $tRNA^{Leu}$ (▽) are plotted. Errors represented are the standard deviation of four independent experiments.

$tRNA^{Phe}$) while the tRNA body contributes to approximately 40-fold difference (comparing Phe-$tRNA^{Phe}$ with rA-Phe) in affinity, further confirming that the tRNA body does contribute significantly to recognition.

5.2.3. Determining the Recognition Element by Hybrid tRNAs

To identify the specific element that is important for L/F transferase tRNA recognition, we synthesized, purified and assayed hybrid tRNAs via "step-by-step" mutations to convert the weak substrate Leu-$tRNA^{Leu}$ (GAG) into the strong substrate Leu-$tRNA^{Leu}$ (CAG). The large differences in the rates of product formation when utilizing these two substrates is apparent in **Figure 5-2B** where the data for these two tRNAs are shown as solid lines. Mutations to the acceptor, D- and T-stems of the tRNA were investigated. As both tRNAs are aminoacylated by leucyl-tRNA synthetase (LeuRS), it was predicted that the mutations would not inhibit aminoacylation. Previous studies have identified the tRNA nucleotides that are essential for *E. coli* LeuRS recognition (Asahara *et al.* 1993a, Asahara *et al.* 1993b, Asahara *et al.* 1998, Larkin *et al.* 2002) and none of the mutations investigated alter these nucleotides. Our assumptions were

validated by performing aminoacylation assays to confirm that these hybrid tRNAs were properly aminoacylated.

5.2.4. tRNA^{Leu} (GAG) Acceptor Stem Hybrids

Figure 5-6A shows the sequences and mutations in the acceptor stem of the tRNA hybrids (constructs 1-11). Mutations in constructs 1-5 focus on the major differences to "step-by-step" convert the weak Leu-tRNA^{Leu} (GAG) substrate into the better Leu-tRNA^{Leu} (CAG) substrate. Construct 1 with the $U_{72}C$ mutation converts the non-canonical G:U to a canonical G:C pair. Construct 2 with the $U_{68}C$ mutation mimics the A:C mismatch in the isoacceptor of Leu-tRNA^{Leu} (CAG). Construct 3 (double mutant) with the G_4A and $C_{69}U$ mutation converts a G:C pair to a weaker A:U pair. Construct 4 (triple mutant) combines the mutations in construct 2 and 3, while construct 5 (quadruple mutant) combines the mutations in construct 1 and 4. Initial rates of product formation determined for all hybrid tRNA constructs are listed in **Table 5-3** and the kinetic parameters are listed in **Table 5-4**. **Figure 5-6B** graphically shows the initial rates of product formation versus tRNA concentration. Subsequent data is

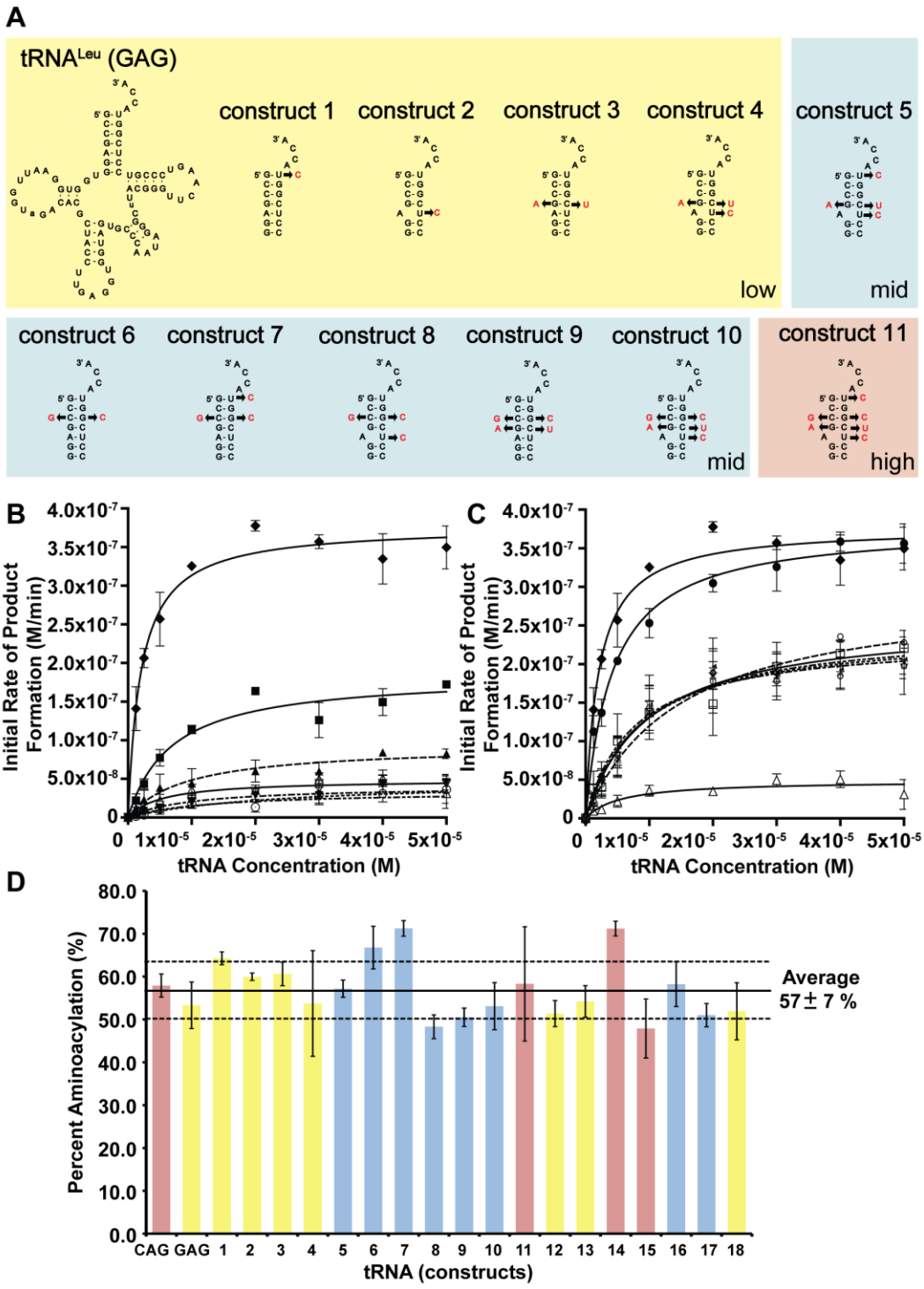
A
tRNA^Leu (GAG)
construct 1
construct 2
construct 3
construct 4
construct 5
low
mid
construct 6
construct 7
construct 8
construct 9
construct 10
construct 11
mid
high
B
C
Initial Rate of Product Formation (M/min)
4.0x10-7
3.5x10-7
3.0x10-7
2.5x10-7
2.0x10-7
1.5x10-7
1.0x10-7
5.0x10-8
0
1x10-5
2x10-5
3x10-5
4x10-5
5x10-5
tRNA Concentration (M)
D
Percent Aminoacylation (%)
80.0
70.0
60.0
50.0
40.0
30.0
20.0
10.0
0.0
CAG GAG 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
tRNA (constructs)
Average
57 ± 7 %

Figure 5-6: Acceptor stem hybrids identify two independent sequence elements for optimal substrate utilization. A) Cloverleaf structures of tRNA hybrid constructs 1-11. **B)** A graphical display of initial rate of product formation versus tRNA concentration for $tRNA^{Leu}$ (CAG) (♦), $tRNA^{Leu}$ (GAG) (Δ), constructs 1 (▲), 2 (▼), 3 (∇), 4 (○), and 5 (■). Errors represented are the standard deviation of three independent experiments. **C)** A graphical display of initial rate of product formation versus tRNA concentration for $tRNA^{Leu}$ (CAG) (♦), $tRNA^{Leu}$ (GAG) (Δ), constructs 6 (□), 7 (♀), 8 (×), 9 (♂), 10 (☆), and 11 (●). Errors represented are the standard deviation of three independent experiments. **D)** A bar graph presenting the maximal percent aminoacylation of all hybrid constructs after 7 minutes of aminoacylation. Errors represented are the standard deviation of three independent experiments.

Table 5-3: Initial reaction rates of L/F transferase catalyzed peptide bond formation by *in vitro* transcribed *E. coli* leucyl-tRNA hybrids.

Construct Number	Name	Mutation	tRNA Conc. (μM)	Initial Reaction Rate ($\mu M\ min^{-1}$)
1	$U_{72}C$	$tRNA^{Leu}$ (GAG) $U_{72}C$	1.3	0.012 ± 0.014
			2.5	0.023 ± 0.021
			5.0	0.039 ± 0.030
			10.0	0.044 ± 0.033
			20.0	0.060 ± 0.025
			30.0	0.060 ± 0.021
			40.0	0.084 ± 0.009
			50.0	0.083 ± 0.011
2	A:C mismatch	$tRNA^{Leu}$ (GAG) $U_{68}C$	1.3	0.003 ± 0.001
			2.5	0.008 ± 0.005
			5.0	0.012 ± 0.009
			10.0	0.019 ± 0.012
			20.0	0.029 ± 0.015
			30.0	0.029 ± 0.018
			40.0	0.044 ± 0.019
			50.0	0.046 ± 0.017
3	Double mutant (DM)	$tRNA^{Leu}$ (GAG) G_4A, $C_{69}U$	1.3	0.003 ± 0.002
			2.5	0.004 ± 0.002
			5.0	0.012 ± 0.006
			10.0	0.023 ± 0.013
			20.0	0.033 ± 0.016
			30.0	0.037 ± 0.021
			40.0	0.044 ± 0.019
			50.0	0.044 ± 0.020
4	Triple mutant (TM)	$tRNA^{Leu}$ (GAG) G_4A, $U_{68}C$, $C_{69}U$	1.3	0.001 ± 0.001
			2.5	0.003 ± 0.002
			5.0	0.011 ± 0.011
			10.0	0.018 ± 0.011
			20.0	0.013 ± 0.011
			30.0	0.029 ± 0.022
			40.0	0.031 ± 0.019
			50.0	0.036 ± 0.031
5	Quadruple mutant	$tRNA^{Leu}$ (GAG) G_4A,	1.3	0.022 ± 0.015
			2.5	0.043 ± 0.012
			5.0	0.077 ± 0.018

	(QM)	$U_{68}C$, $C_{69}U$, $U_{72}C$	10.0 20.0 30.0 40.0 50.0	0.114 ± 0.009 0.164 ± 0.009 0.126 ± 0.039 0.149 ± 0.030 0.172 ± 0.002
6	C:G swap	$tRNA^{Leu}$ (GAG) C_3G, $G_{70}C$	1.3 2.5 5.0 10.0 20.0 30.0 40.0 50.0	0.038 ± 0.040 0.041 ± 0.024 0.100 ± 0.061 0.137 ± 0.060 0.148 ± 0.070 0.196 ± 0.033 0.214 ± 0.028 0.221 ± 0.025
7	C:G swap + $U_{72}C$	$tRNA^{Leu}$ (GAG) C_3G, $G_{70}C$, $U_{72}C$	1.3 2.5 5.0 10.0 20.0 30.0 40.0 50.0	0.019 ± 0.014 0.051 ± 0.038 0.078 ± 0.044 0.139 ± 0.050 0.174 ± 0.051 0.186 ± 0.047 0.232 ± 0.005 0.224 ± 0.009
8	C:G swap + A:C mismatch	$tRNA^{Leu}$ (GAG) C_3G, $U_{68}C$, $G_{70}C$	1.3 2.5 5.0 10.0 20.0 30.0 40.0 50.0	0.032 ± 0.023 0.049 ± 0.023 0.080 ± 0.044 0.137 ± 0.061 0.185 ± 0.084 0.183 ± 0.051 0.199 ± 0.056 0.202 ± 0.072
9	C:G swap + DM	$tRNA^{Leu}$ (GAG) C_3G, G_4A, $C_{69}U$, $G_{70}C$	1.3 2.5 5.0 10.0 20.0 30.0 40.0 50.0	0.032 ± 0.020 0.048 ± 0.029 0.088 ± 0.033 0.143 ± 0.016 0.192 ± 0.043 0.181 ± 0.002 0.188 ± 0.004 0.201 ± 0.025
10	C:G swap + TM	$tRNA^{Leu}$ (GAG) C_3G,	1.3 2.5 5.0	0.024 ± 0.013 0.041 ± 0.022 0.071 ± 0.026

		G_4A,	10.0	0.150 ± 0.062
		$U_{68}C$,	20.0	0.194 ± 0.045
		$C_{69}U$,	30.0	0.200 ± 0.049
		$G_{70}C$	40.0	0.189 ± 0.035
			50.0	0.198 ± 0.030
11	C:G swap + QM = full CAG accept or stem	$tRNA^{Leu}$ (GAG) C_3G, G_4A, $U_{68}C$, $C_{69}U$, $G_{70}C$, $U_{72}C$	1.3	0.112 ± 0.036
			2.5	0.137 ± 0.030
			5.0	0.204 ± 0.007
			10.0	0.253 ± 0.033
			20.0	0.305 ± 0.020
			30.0	0.326 ± 0.054
			40.0	0.358 ± 0.021
			50.0	0.356 ± 0.044
12	D-stem	$tRNA^{Leu}$ (GAG) $U_{11}C$, $A_{24}G$	1.3	0.008 ± 0.001
			2.5	0.014 ± 0.007
			5.0	0.037 ± 0.011
			10.0	0.056 ± 0.015
			20.0	0.058 ± 0.017
			30.0	0.050 ± 0.014
			40.0	0.063 ± 0.031
			50.0	0.062 ± 0.027
13	T-stem	$tRNA^{Leu}$ (GAG) $A_{49}G$, $U_{65}C$	1.3	0.010 ± 0.004
			2.5	0.027 ± 0.004
			5.0	0.040 ± 0.004
			10.0	0.057 ± 0.008
			20.0	0.064 ± 0.003
			30.0	0.062 ± 0.011
			40.0	0.068 ± 0.010
			50.0	0.077 ± 0.007
14	C:G swap + QM + D-stem	$tRNA^{Leu}$ (GAG) C_3G, G_4A, $U_{11}C$, $A_{24}G$, $U_{68}C$, $C_{69}U$, $G_{70}C$, $U_{72}C$	1.3	0.071 ± 0.012
			2.5	0.143 ± 0.082
			5.0	0.244 ± 0.077
			10.0	0.277 ± 0.078
			20.0	0.321 ± 0.046
			30.0	0.329 ± 0.044
			40.0	0.334 ± 0.039
			50.0	0.372 ± 0.040
15	C:G	$tRNA^{Leu}$	1.3	0.089 ± 0.025

	swap + QM + D-stem + T-stem	(GAG) C_3G, G_4A, $U_{11}C$, $A_{24}G$, A49G, $U_{65}C$, $U_{68}C$, $C_{69}U$, $G_{70}C$, $U_{72}C$	2.5 5.0 10.0 20.0 30.0 40.0 50.0	0.134 ± 0.008 0.213 ± 0.048 0.239 ± 0.031 0.289 ± 0.043 0.313 ± 0.061 0.337 ± 0.070 0.368 ± 0.055
16	Reverse C:G swap	$tRNA^{Leu}$ (CAG) G_3C, $C_{70}G$	1.3 2.5 5.0 10.0 20.0 30.0 40.0 50.0	0.025 ± 0.014 0.048 ± 0.025 0.080 ± 0.047 0.146 ± 0.066 0.170 ± 0.057 0.189 ± 0.023 0.198 ± 0.029 0.198 ± 0.027
17	Reverse QM	$tRNA^{Leu}$ (CAG) A_4G, $C_{68}U$, $U_{69}C$, $C_{72}U$	1.3 2.5 5.0 10.0 20.0 30.0 40.0 50.0	0.022 ± 0.014 0.052 ± 0.027 0.076 ± 0.036 0.162 ± 0.037 0.186 ± 0.038 0.191 ± 0.042 0.209 ± 0.046 0.213 ± 0.051
18	Reverse C:G swap + QM = full acceptor stem of GAG	$tRNA^{Leu}$ (CAG) G_3C, A_4G, $C_{68}U$, $U_{69}C$, $C_{70}G$, $C_{72}U$	1.3 2.5 5.0 10.0 20.0 30.0 40.0 50.0	0.008 ± 0.006 0.022 ± 0.016 0.028 ± 0.021 0.060 ± 0.012 0.071 ± 0.005 0.086 ± 0.028 0.104 ± 0.018 0.100 ± 0.021

Errors represented are the standard deviation of three independent experiments. Nucleotides numbering is according to (Sprinzl *et al.* 1998).

Table 5-4: Kinetic parameters of L/F transferase catalyzed peptide bond formation by *in vitro* transcribed leucyl-tRNA hybrids.

Construct Number	Name	Mutation	Apparent K_M (μM)	Apparent k_{cat} (min^{-1})	Catalytic Efficiency (k_{cat}/K_M / k_{cat}^{CAG}/K_M^{CAG})
1	$U_{72}C$	$tRNA^{Leu}$ (GAG) $U_{72}C$	9.3 ± 4.3	0.025 ± 0.004	0.054
2	A:C mismatch	$tRNA^{Leu}$ (GAG) $U_{68}C$	24.2 ± 12.2	0.013 ± 0.003	0.011
3	Double mutant (DM)	$tRNA^{Leu}$ (GAG) G_4A, $C_{69}U$	9.9 ± 4.6	0.011 ± 0.002	0.022
4	Triple mutant (TM)	$tRNA^{Leu}$ (GAG) G_4A, $U_{68}C$, $C_{69}U$	12.9 ± 11.7	0.009 ± 0.003	0.014
5	Quadruple mutant (QM)	$tRNA^{Leu}$ (GAG) G_4A, $U_{68}C$, $C_{69}U$, $U_{72}C$	6.9 ± 1.7	0.049 ± 0.003	0.142
6	C:G swap	$tRNA^{Leu}$ (GAG) C_3G, $G_{70}C$	9.8 ± 3.4	0.069 ± 0.008	0.141
7	C:G swap + $U_{72}C$	$tRNA^{Leu}$ (GAG) C_3G, $G_{70}C$, $U_{72}C$	14.7 ± 4.1	0.079 ± 0.008	0.107
8	C:G swap + A:C mismatch	$tRNA^{Leu}$ (GAG) C_3G, $U_{68}C$, $G_{70}C$	8.6 ± 3.6	0.065 ± 0.008	0.151
9	C:G swap + DM	$tRNA^{Leu}$ (GAG) C_3G, G_4A, $C_{69}U$, $G_{70}C$	7.4 ± 1.5	0.062 ± 0.004	0.168
10	C:G swap + TM	$tRNA^{Leu}$ (GAG) C_3G, G_4A, $U_{68}C$, $C_{69}U$, $G_{70}C$	8.8 ± 2.8	0.066 ± 0.006	0.150
11	C:G swap + QM = full CAG acceptor	$tRNA^{Leu}$ (GAG) C_3G, G_4A, $U_{68}C$, $C_{69}U$, $G_{70}C$,	4.2 ± 0.6	0.101 ± 0.004	0.481

	stem	$U_{72}C$			
12	D-stem	$tRNA^{Leu}$ (GAG) $U_{11}C$, $A_{24}G$	7.4 ± 2.4	0.022 ± 0.002	0.059
13	T-stem	$tRNA^{Leu}$ (GAG) $A_{49}G$, $U_{65}C$	5.2 ± 0.8	0.022 ± 0.001	0.085
14	C:G swap + QM + D-stem	$tRNA^{Leu}$ (GAG) C_3G, G_4A, $U_{11}C$, $A_{24}G$, $U_{68}C$, $C_{69}U$, $G_{70}C$, $U_{72}C$	3.9 ± 0.9	0.102 ± 0.006	0.523
15	C:G swap + QM + D-stem + T-stem	$tRNA^{Leu}$ (GAG) C_3G, G_4A, $U_{11}C$, $A_{24}G$, A49G, $U_{65}C$, $U_{68}C$, $C_{69}U$, $G_{70}C$, $U_{72}C$	4.5 ± 0.9	0.099 ± 0.005	0.440
16	Reverse C:G swap	$tRNA^{Leu}$ (CAG) G_3C, $C_{70}G$	8.5 ± 2.6	0.064 ± 0.006	0.151
17	Reverse QM	$tRNA^{Leu}$ (CAG) A_4G, $C_{68}U$, $U_{69}C$, $C_{72}U$	8.7 ± 2.5	0.068 ± 0.006	0.156
18	Reverse C:G swap + QM = full acceptor stem of GAG	$tRNA^{Leu}$ (CAG) G_3C, A_4G, $C_{68}U$, $U_{69}C$, $C_{70}G$, $C_{72}U$	15.1 ± 4.8	0.035 ± 0.004	0.046

Errors represented are the standard deviation of three independent experiments.

Nucleotides numbering is according to (Sprinzl *et al.* 1998).

compared to the wild-type GAG isoacceptor as reference. Although constructs 1-4 feature the major acceptor stem sequence differences, they remain to be poor L/F transferase substrates (1.7 to 4.5-fold increase in apparent K_M compared to GAG, and lower catalytic efficiency). Interestingly, construct 5 (quadruple mutant) exhibits an enhanced utilization by L/F transferase activity to a midpoint to that of the optimal substrate CAG when the four nucleotide mutations combined.

Figure 5-6C graphically shows the initial rates of product formation versus tRNA concentration for constructs 6-11 in converting isoacceptor GAG to CAG. Construct 6 (C:G swap) swaps the C:G pair of GAG isoacceptor to a G:C pair of CAG isoacceptor, a relative conserved modification in the acceptor stem. Surprisingly, this conserved C:G swap in the acceptor stem alone enhanced L/F transferase activity to a midpoint level when compare to the optimal substrate CAG. Next, we wanted to determine whether we could improve the activity of GAG with minimal modifications to its acceptor stem. We combined the C:G swap mutation with the mutations in constructs 1-4, and named those constructs 7-10. Kinetic analysis show that C:G swap in combination with $U_{72}C$, $U_{68}C$,

double mutant, or the triple mutant do not have any improvement on substrate utilization. Not until construct 11 (C:G swap + quadruple mutant), essentially the full acceptor stem of CAG, increases the apparent K_M slightly but improves the apparent k_{cat} to wild-type CAG levels. A total of 10-fold improvement in catalytic efficiency is achieved with mutations in the acceptor stem of GAG alone.

The maximal percent aminoacylation after 7 minutes for each tRNA hybrid constructs are shown in a bar graph in **Figure 5-6D** (for full time course see **Figure 5-7**). Again, there are no significant differences between the aminoacylation of these tRNA hybrids compared to the wild-type isoacceptors. The changes in aminoacylation of the hybrid constructs are not sufficient for changes in the initial reaction rates.

Here we have demonstrated that mutations in the GAG tRNA acceptor stem alone are able to optimize the utilization of aa-tRNAs by L/F transferase to the maximal CAG isoacceptor levels. We have determined two independent sequence elements – the C:G swap at position 3:70 and the quadruple mutations (G_4A, $U_{68}C$, $C_{69}U$, $U_{72}C$) – within the acceptor stem that are important for L/F transferase recognition and catalysis.

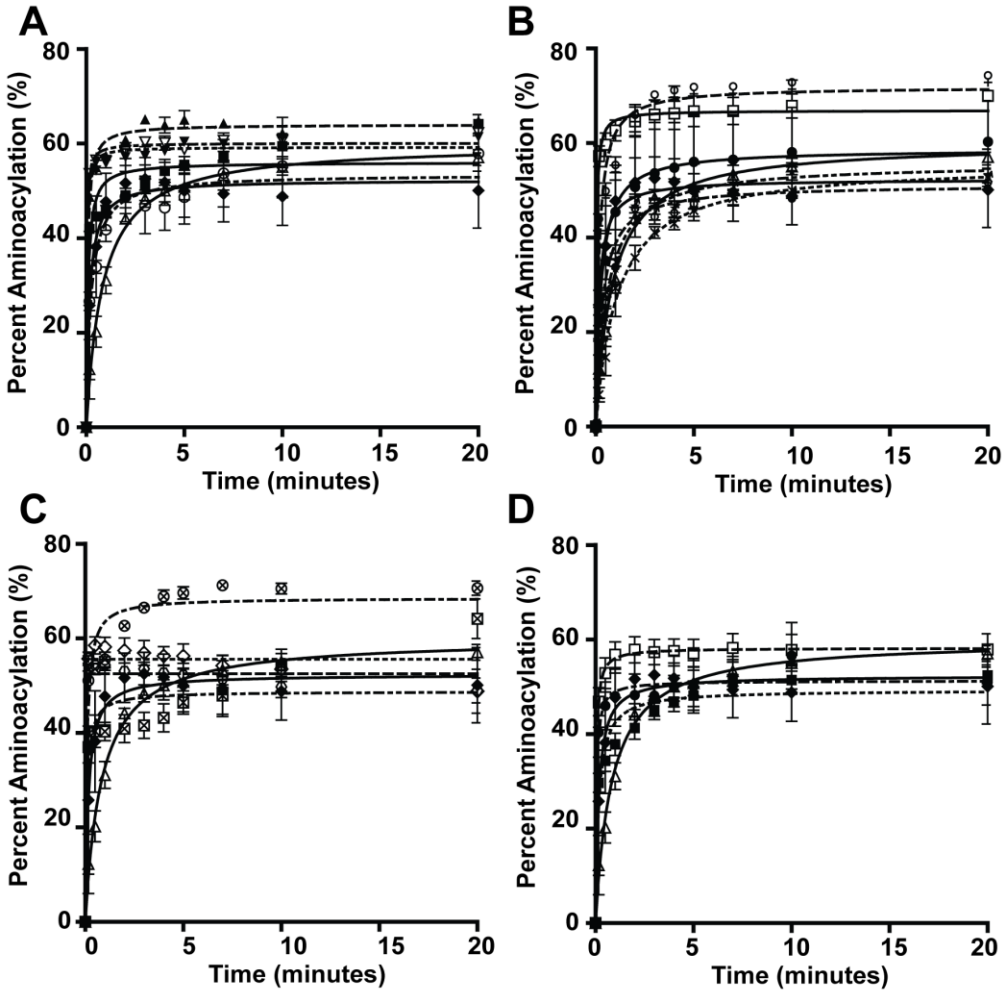

Figure 5-7: Aminoacylation time courses show that the hybrid constructs are efficiently aminoacylated by LeuRS. **A)** Percent aminoacylation versus time graph for $tRNA^{Leu}$ (CAG) (♦), $tRNA^{Leu}$ (GAG) (Δ), constructs 1 (▲), 2 (▼), 3 (∇), 4 (○), and 5 (■). **B)** Percent aminoacylation versus time graph for $tRNA^{Leu}$ (CAG) (♦), $tRNA^{Leu}$ (GAG) (Δ), constructs 6 (□), 7 (♀), 8 (×), 9 (♂), 10 (☆), and 11 (●). **C)** Percent aminoacylation versus time graph for $tRNA^{Leu}$ (CAG) (♦), $tRNA^{Leu}$ (GAG) (Δ), constructs 12 (⎔), 13 (◊), 14 (⊗), and 15 (⊠). **D)** Percent aminoacylation versus time graph for $tRNA^{Leu}$ (CAG) (♦), $tRNA^{Leu}$ (GAG) (Δ), constructs 16 (□), 17 (■), and 18 (●). Errors represented are the standard deviation of three independent experiments.

5.2.5. $tRNA^{Leu}$ (GAG) D- and T-stem Hybrids

To further determine whether other parts of the tRNA body contributes to L/F transferase recognition and catalysis, we generated constructs 12-15 with mutations focused on the D- and T-stem of Leu-$tRNA^{Leu}$ (GAG) (**Figure 5-8A**). Construct 12 (D-stem) with the mutations $U_{11}C$ and $A_{24}G$ converts the entire D-stem/D-loop to mimic CAG's. Construct 13 (T-stem) with the mutations $A_{49}G$ and $U_{65}C$ converts the weaker A:U pair to a stronger G:C pair in the T-stem. Construct 14 (C:G swap + QM + D-stem) and construct 15 (C:G swap + QM + D-stem + T-stem) combines the full acceptor stem mutations with D- and T-stem mutations. **Figure 5-8B** clearly shows that D- or T-stem mutations alone do not significantly alter the utilization of the aa-tRNA by L/F transferase. Additionally, full acceptor stem mutations in combination with D-stem and T-stem mutations also do not significantly alter the utilization of construct 11. This suggests that D- and T-stem do not play a significant role in L/F transferase recognition and catalysis.

5.2.6. $tRNA^{Leu}$ (CAG) Reverse Hybrids

To validate our findings on the two independent sequence elements recognized by L/F transferase, we generated reverse hybrids to convert the optimal CAG substrate into the poorer GAG substrate. The sequence and mutations are depicted in **Figure 5-9A**. Construct 16 (reverse C:G swap) is a reversal of the C:G swap hybrid, meanwhile construct 17 (reverse quadruple mutant) with the mutation A_4G, $C_{68}U$, $U_{69}C$, and $C_{72}U$

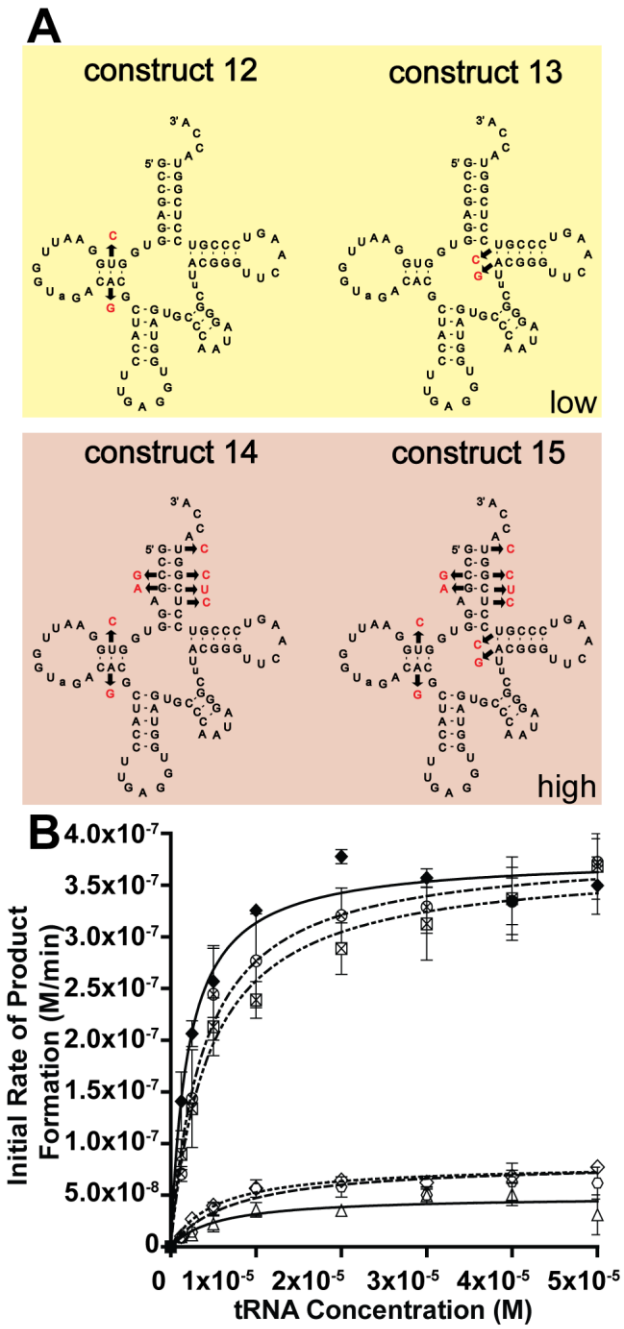

Figure 5-8: No significant recognition contribution by the D-stem and T-stem of the tRNA body. A) Cloverleaf structures of tRNA hybrid constructs 12-15. **B)** A graphical display of initial rate of product formation versus tRNA concentration for $tRNA^{Leu}$ (CAG) (♦), $tRNA^{Leu}$ (GAG) (Δ), constructs 12 (⎔), 13 (◊), 14 (⊗), and 15 (⊠). Errors represented are the standard deviation of three independent experiments.

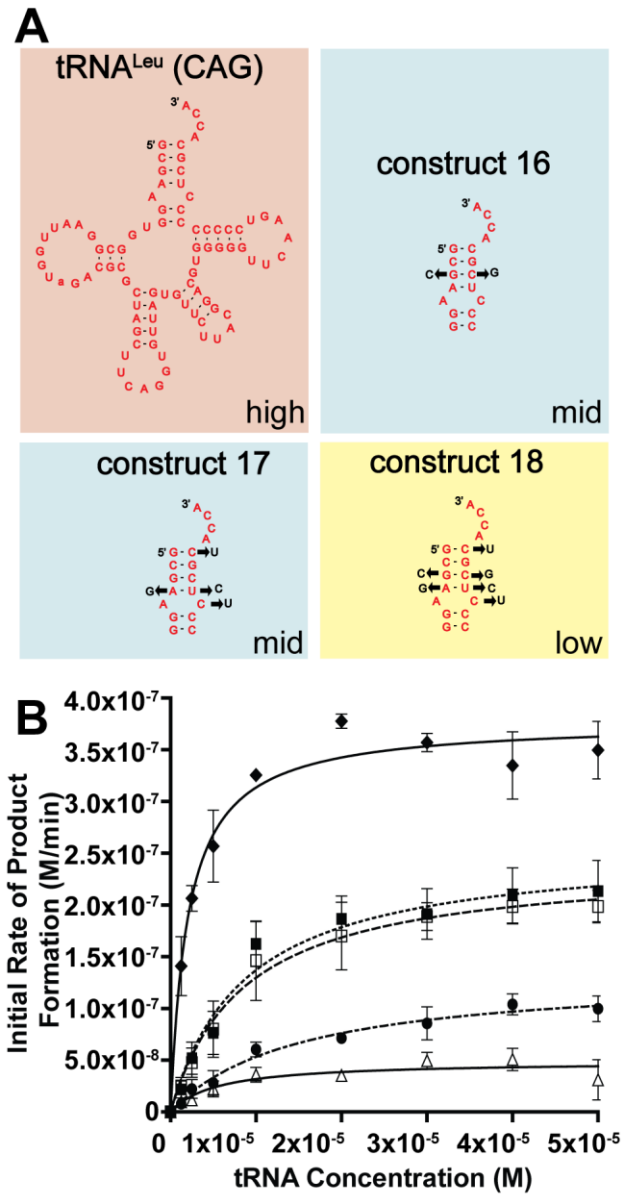

Figure 5-9: Reverse hybrids validate the identified two independent sequence elements for optimal substrate utilization. A) Cloverleaf structures of tRNA hybrid construct 16-18. **B)** A graphical display of initial rate of product formation versus tRNA concentration for $tRNA^{Leu}$ (CAG) (♦), $tRNA^{Leu}$ (GAG) (Δ), constructs 16 (□), 17 (■), and 18 (●). Errors represented are the standard deviation of three independent experiments.

is a reversal of the quadruple mutant. Construct 18 (reverse C:G + QM) combines the mutations in constructs 16 and 17 to generate the full acceptor stem of GAG isoacceptor in the context of CAG isoacceptor. **Figure 5-9B** shows the graph of initial rates of product formation versus tRNA concentration. Subsequent data is compared to the wild-type CAG isoacceptor as reference. The reverse C:G swap and reverse quadruple mutant, as predicted, independently decrease the affinity (apparent K_M increase by 4.3-fold compared to CAG) and decrease the relative catalytic efficiency of the CAG isoacceptor to a midway level (6.4 to 6.6-fold decrease in catalytic efficiency). The reverse full acceptor stem of GAG (reverse C:G swap + QM) significantly decrease the apparent K_M by 7.5-fold and decrease the relative catalytic efficiency to a lower level similar to that of the GAG isoacceptor (21.7-fold decrease in catalytic efficiency compared to CAG). These data confirms the importance of the C:G swap and quadruple mutant in L/F transferase recognition and catalysis.

5.3. Discussion

5.3.1. The Acceptor Stem of an aa-tRNA is Important for L/F transferase Recognition

An atypical function for tRNA is their role in tRNA-dependent post-translational addition of amino acids to the N-terminus of proteins, which leads to protein degradation (Bachmair and Varshavsky 1989). The eubacterial L/F transferase catalyzes the transfer of a Leu or Phe (or to a lesser extend Met) from a cognate aa-tRNA onto the N-terminus of a

protein polypeptide (Leibowitz and Soffer 1969, Scarpulla *et al.* 1976). The protein peptide substrate specificity has been well studied with the aid of X-ray crystal structures and *in vitro* enzymatic assays (Mogk *et al.* 2007, Watanabe *et al.* 2007, Wang *et al.* 2008a, Wang *et al.* 2008b, Ninnis *et al.* 2009, Schuenemann *et al.* 2009, Kawaguchi *et al.* 2013). Nonetheless, the aa-tRNA recognition by L/F transferase has remained somewhat elusive. Although there are tRNA substrate analogues (3' rA-Phe and puromycin) bound X-ray crystal structures, the molecular insights remained within the 3' aminoacyl adenosine as an intact aa-tRNA: protein structure has not been solved (Suto *et al.* 2006, Watanabe *et al.* 2007).

Here we investigated L/F transferase's preference for a specific $tRNA^{Leu}$ isoacceptor, and subsequently determine the nucleotides of an aa-tRNA that are optimal for substrate utilization. Our results indicated that the $tRNA^{Leu}$ (CAG) isoacceptor is the optimal L/F transferase substrate. Using *in vitro* transcribed hybrid tRNAs, we identified two independent sequence elements in this optimal tRNA substrate including the G_3:C_{70} base pair and a set of four nucleotides (C_{72}, A_4:U_{69}, C_{68}) at the acceptor stem that is shown to be important for binding and catalysis. Our data does not support the weak acceptor stem hypothesis proposed (Abramochkin and Shrader 1996), since individual mutants that generate weak base pairs (i.e. A:U or A:C mismatch in constructs 2 and 3) have no significant effects on activity while a single conserved C:G or G:C swap at position 3:70 (in constructs 6 and 16) significantly modifies L/F transferase

activity. Comparing the 3:70 base pair on wild type tRNA isoacceptors (**Figure 5-1**), we observed that the two high activity Leu-tRNA^{Leu}s CAG and UAG contain the G_3:C_{70} base pair while the remaining tRNAs contain the C_3:G_{70} base pair. Thus, the G_3:C_{70} base pair may serve as a simple predictor for tRNA substrate utilization by L/F transferase. Additionally, the set of four nucleotides only enhances binding and catalysis when in combination (construct 5) but individual mutation effects (constructs 1-4) are relatively insignificant. The G_1:C_{72} base pair is common between the CAG, UAG, and CAA isoacceptors, the three better tRNA^{Leu} isoacceptors of the five, whereas the A_4:U_{69} base pair and A_5:C_{68} mismatch are unique features of CAG. We hypothesize that the sequence elements may function in combination by contributing to the overall helical shape of the acceptor stem for efficient substrate recognition and catalysis. Future experiments with chemical acylation of various amino acids to various acceptor stem helices may provide additional insights into determining the relative contribution of the amino acid and tRNA to L/F transferase binding affinity and thus the molecular mechanism of aa-tRNA recognition.

5.3.2. A Proposed Model of L/F transferase aa-tRNA Recognition

Our investigations have added to the current model of aa-tRNA recognition by L/F transferase that has been developed from previous biochemical and structural studies (Leibowitz and Soffer 1971, Abramochkin and Shrader 1996, Suto *et al.* 2006, Watanabe *et al.* 2007). L/F transferase, like other aa-tRNA binding enzymes, recognizes both the

esterified amino acid as well as sequence-specific elements within the tRNA body for recognition (Asahara *et al.* 1993a, Banerjee *et al.* 2010, Schrader *et al.* 2011). The amino acid selectivity is mainly through the C-shaped hydrophobic pocket of L/F transferase, which sterically prevents larger β-branched amino acids (i.e. Ile and Val) and disfavours smaller amino acids (i.e. Ala and Pro) as they are not large enough to make sufficient hydrophobic contacts (Suto *et al.* 2006, Watanabe *et al.* 2007). **Figure 5-10** shows a proposed docking model of aa-tRNA recognition by L/F transferase. Based on the close proximity of the positive cluster (R76, R80, K83, and R84) of L/F transferase to the acceptor stem of an aa-tRNA, we suggest that the positive cluster may contribute to the specific recognition of the acceptor stem. Thus, the recognition of the tRNA body involves the 3' aminoacyl adenosine, the major determinant G_3:C_{70} base pair, the combined set of four nucleotides (C_{72}, A_4:U_{69}, C_{68}), and the sequence independent recognition of the D-stem. We hypothesize that the eukaryotic aminoacyl-tRNA protein transferase (ATE1) may similarly depend on both the esterified amino acid and sequence-specific determinants on the tRNA body (i.e. likely the acceptor stem) for efficient aa-tRNA recognition.

5.3.3. Codon Usage, Abundance, and Aminoacylation Efficiency of $tRNA^{Leu}$ Isoacceptors

As the $tRNA^{Leu}$ (CAG) isoacceptor was determined to be the optimal substrate, we evaluated the literature regarding this isoacceptor

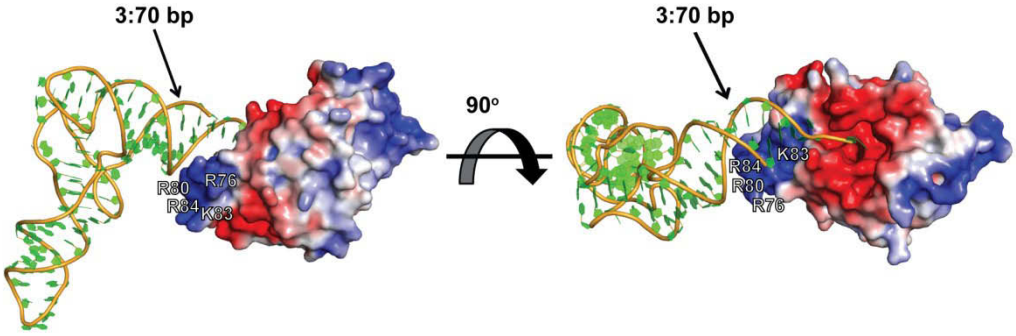

Figure 5-10: A proposed docking model of aminoacyl-tRNA binding to L/F transferase. Our model suggests that in addition to the 3' aminoacyl adenosine recognition and electrostatic interaction, the positive cluster (R76, R80, K83, and R84) of L/F transferase may play a role in the specific recognition of the acceptor stem of an aminoacyl-tRNA. To generate the model, the structure of L/F transferase-rA-Phe complex (shown as electrostatic surface, PDB ID: 2Z3K) (Watanabe *et al.* 2007) was superimposed to the FemX-peptidyl-RNA complex (PDB ID: 4II9) (Fonvielle *et al.* 2013) via the conserved core of the GNAT domain. The combined 3' CCA end (the C_{74} and C_{75} of the peptidyl-RNA and adenosine of rA-Phe) were then used as references to dock the yeast $tRNA^{Phe}$ (shown as ribbon, PDB ID: 1EHZ) (Shi and Moore 2000). The model was generated using PyMOL (version 1.41) and electrostatic potentials were calculated by APBS (version 1.8).

and what has been reported with respect to abundance, codon bias and amino acid dependent aminoacylation changes. $tRNA^{Leu}$ (CAG) is the most abundant leucine isoacceptor in *E. coli* representing 50% of all $tRNA^{Leu}$ isoacceptors (Dong *et al.* 1996) and decodes the most frequently used 5'-CUG-3' codon across various growth rates (Emilsson and Kurland 1990, Dong *et al.* 1996) and media conditions (Holmes *et al.* 1977). The CAG isoacceptor is therefore widely used during protein synthesis and do not appear to be idiosyncratic for the post-translational addition of amino acids. However during leucine starvation, $tRNA^{Leu}$ (CAG) aminoacylation level rapidly decreases to 9% of its steady-state levels yet it also increases rapidly upon restoration of leucine levels (Elf *et al.* 2003, Dittmar *et al.* 2005, Sorensen *et al.* 2005). It has been suggested that this differential aminoacylation may serve as a quick response to environmental stress (Elf *et al.* 2003, Dittmar *et al.* 2005, Sorensen *et al.* 2005). These rapid changes in $tRNA^{Leu}$ (CAG) aminoacylation levels during environmental stress may allow L/F transferase to compete with EF-Tu for aa-tRNA as substrates.

5.3.4. Concluding Remarks

In conclusion, we have demonstrated that the most abundant leucyl-tRNA in *E. coli*, $tRNA^{Leu}$ (CAG), is the most optimal substrate for L/F transferase. We confirmed that the rate differences are not due to differential aminoacylation. Using “step-by-step” hybrid tRNAs, we have identified two independent sequence elements on the acceptor stem of

Leu-tRNA^{Leu} (CAG) that are important for optimal binding and catalysis by L/F transferase. A G_3:C_{70} base pair and a set of four nucleotides in combination (C_{72}, A_4:U_{69}, C_{68}) contribute to optimal tRNA recognition by L/F transferase. This maps a more specific, sequence-dependent tRNA recognition model of L/F transferase than previously thought.

5.4. References

Abramochkin, G., and Shrader, T.E. (1995) The leucyl/phenylalanyl-tRNA-protein transferase. Overexpression and characterization of substrate recognition, domain structure, and secondary structure. *J.Biol.Chem.* **270,** 20621-20628

Abramochkin, G., and Shrader, T.E. (1996) Aminoacyl-tRNA recognition by the leucyl/phenylalanyl-tRNA-protein transferase. *J.Biol.Chem.* **271,** 22901-22907

Agirrezabala, X., and Frank, J. (2009) Elongation in translation as a dynamic interaction among the ribosome, tRNA, and elongation factors EF-G and EF-Tu. *Q.Rev.Biophys.* **42,** 159-200

Andersen, C., and Wiborg, O. (1994) *Escherichia coli* elongation-factor-Tu mutants with decreased affinity for aminoacyl-tRNA. *Eur.J.Biochem.* **220,** 739-744

Asahara, H., Himeno, H., Tamura, K., Hasegawa, T., Watanabe, K., and Shimizu, M. (1993a) Recognition nucleotides of *Escherichia coli* tRNA^{Leu} and its elements facilitating discrimination from tRNA^{Ser} and tRNA^{Tyr}. *J.Mol.Biol.* **231,** 219-229

Asahara, H., Himeno, H., Tamura, K., Nameki, N., Hasegawa, T., and Shimizu, M. (1993b) Discrimination among *E. coli* tRNAs with a long variable arm. *Nucleic Acids Symp.Ser.* **(29),** 207-208

Asahara, H., Nameki, N., and Hasegawa, T. (1998) *In vitro* selection of RNAs aminoacylated by *Escherichia coli* leucyl-tRNA synthetase. *J.Mol.Biol.* **283,** 605-618

Bachmair, A., and Varshavsky, A. (1989) The degradation signal in a short-lived protein. *Cell.* **56,** 1019-1032

Bailly, M., Blaise, M., Lorber, B., Becker, H.D., and Kern, D. (2007) The transamidosome: a dynamic ribonucleoprotein particle dedicated to prokaryotic tRNA-dependent asparagine biosynthesis. *Mol.Cell.* **28,** 228-239

Banerjee, R., Chen, S., Dare, K., Gilreath, M., Praetorius-Ibba, M., Raina, M., Reynolds, N.M., Rogers, T., Roy, H., Yadavalli, S.S., and Ibba, M. (2010) tRNAs: cellular barcodes for amino acids. *FEBS Lett.* **584,** 387-395

Becker, H.D., and Kern, D. (1998) *Thermus thermophilus*: a link in evolution of the tRNA-dependent amino acid amidation pathways. *Proc.Natl.Acad.Sci.U.S.A.* **95,** 12832-12837

Chan, P.P., and Lowe, T.M. (2009) GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res.* **37,** D93-7

Dittmar, K.A., Sorensen, M.A., Elf, J., Ehrenberg, M., and Pan, T. (2005) Selective charging of tRNA isoacceptors induced by amino-acid starvation. *EMBO Rep.* **6,** 151-157

Dong, H., Nilsson, L., and Kurland, C.G. (1996) Co-variation of tRNA abundance and codon usage in *Escherichia coli* at different growth rates. *J.Mol.Biol.* **260,** 649-663

Ebhardt, H.A., Xu, Z., Fung, A.W., and Fahlman, R.P. (2009) Quantification of the post-translational addition of amino acids to proteins by MALDI-TOF mass spectrometry. *Anal.Chem.* **81,** 1937-1943

Elf, J., Nilsson, D., Tenson, T., and Ehrenberg, M. (2003) Selective charging of tRNA isoacceptors explains patterns of codon usage. *Science.* **300,** 1718-1722

Emilsson, V., and Kurland, C.G. (1990) Growth rate dependence of transfer RNA abundance in *Escherichia coli. EMBO J.* **9,** 4359-4366

Fonvielle, M., Chemama, M., Villet, R., Lecerf, M., Bouhss, A., Valery, J.M., Etheve-Quelquejeu, M., and Arthur, M. (2009) Aminoacyl-tRNA recognition by the FemXWv transferase for bacterial cell wall synthesis. *Nucleic Acids Res.* **37,** 1589-1601

Francklyn, C.S., and Minajigi, A. (2010) tRNA as an active chemical scaffold for diverse chemical transformations. *FEBS Lett.* **584,** 366-375

Fung, A.W., Ebhardt, H.A., Abeysundara, H., Moore, J., Xu, Z., and Fahlman, R.P. (2011) An alternative mechanism for the catalysis of

peptide bond formation by L/F transferase: substrate binding and orientation. *J.Mol.Biol.* **409,** 617-629

Fung, A.W., Ebhardt, H.A., Krishnakumar, K.S., Moore, J., Xu, Z., Strazewski, P., and Fahlman, R.P. (2014) Probing the Leucyl/Phenylalanyl tRNA Protein Transferase Active Site with tRNA Substrate Analogues. *Protein Pept.Lett.* **21,** 603-614

Giannouli, S., Kyritsis, A., Malissovas, N., Becker, H.D., and Stathopoulos, C. (2009) On the role of an unusual $tRNA^{Gly}$ isoacceptor in *Staphylococcus aureus. Biochimie.* **91,** 344-351

Holmes, W.M., Goldman, E., Miner, T.A., and Hatfield, G.W. (1977) Differential utilization of leucyl-tRNAs by *Escherichia coli. Proc.Natl.Acad.Sci.U.S.A.* **74,** 1393-1397

Ibba, M., and Soll, D. (2004) Aminoacyl-tRNAs: setting the limits of the genetic code. *Genes Dev.* **18,** 731-738

Kawaguchi, J., Maejima, K., Kuroiwa, H., and Taki, M. (2013) Kinetic analysis of the leucyl/phenylalanyl-tRNA-protein transferase with acceptor peptides possessing different N-terminal penultimate residues. *FEBS Open Bio.* **3,** 252-255

LaRiviere, F.J., Wolfson, A.D., and Uhlenbeck, O.C. (2001) Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. *Science.* **294,** 165-168

Larkin, D.C., Williams, A.M., Martinis, S.A., and Fox, G.E. (2002) Identification of essential domains for *Escherichia coli* tRNA(leu) aminoacylation and amino acid editing using minimalist RNA molecules. *Nucleic Acids Res.* **30,** 2103-2113

Leibowitz, M.J., and Soffer, R.L. (1969) A soluble enzyme from *Escherichia coli* which catalyzes the transfer of leucine and phenylalanine from tRNA to acceptor proteins. *Biochem.Biophys.Res.Commun.* **36,** 47-53

Leibowitz, M.J., and Soffer, R.L. (1971) Enzymatic modification of proteins. VII. Substrate specificity of leucyl,phenylalanyl-transfer ribonucleic acid-protein transferase. *J.Biol.Chem.* **246,** 5207-5212

Marshall, R.A., Aitken, C.E., Dorywalska, M., and Puglisi, J.D. (2008) Translation at the single-molecule level. *Annu.Rev.Biochem.* **77,** 177-203

Mogk, A., Schmidt, R., and Bukau, B. (2007) The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies. *Trends Cell Biol.* **17,** 165-172

Ninnis, R.L., Spall, S.K., Talbo, G.H., Truscott, K.N., and Dougan, D.A. (2009) Modification of PATase by L/F-transferase generates a ClpS-dependent N-end rule substrate in *Escherichia coli. EMBO J.* **28,** 1732-1744

Nolan, E.M., and Walsh, C.T. (2009) How nature morphs peptide scaffolds into antibiotics. *Chembiochem.* **10,** 34-53

Rao, P.M., and Kaji, H. (1974) Utilization of isoaccepting leucyl-tRNA in the soluble incorporation system and protein synthesizing systems from *E.coli. FEBS Lett.* **43,** 199-202

Roy, H., and Ibba, M. (2008) RNA-dependent lipid remodeling by bacterial multiple peptide resistance factors. *Proc.Natl.Acad.Sci.U.S.A.* **105,** 4667-4672

Scarpulla, R.C., Deutch, C.E., and Soffer, R.L. (1976) Transfer of methionyl residues by leucyl, phenylalanyl-tRNA-protein transferase. *Biochem.Biophys.Res.Commun.* **71,** 584-589

Schmeing, T.M., and Ramakrishnan, V. (2009) What recent ribosome structures have revealed about the mechanism of translation. *Nature.* **461,** 1234-1242

Schrader, J.M., Chapman, S.J., and Uhlenbeck, O.C. (2011) Tuning the affinity of aminoacyl-tRNA to elongation factor Tu for optimal decoding. *Proc.Natl.Acad.Sci.U.S.A.* **108,** 5215-5220

Schuenemann, V.J., Kralik, S.M., Albrecht, R., Spall, S.K., Truscott, K.N., Dougan, D.A., and Zeth, K. (2009) Structural basis of N-end rule substrate recognition in *Escherichia coli* by the ClpAP adaptor protein ClpS. *EMBO Rep.* **10,** 508-514

Sheppard, K., Yuan, J., Hohn, M.J., Jester, B., Devine, K.M., and Soll, D. (2008) From one amino acid to another: tRNA-dependent amino acid biosynthesis. *Nucleic Acids Res.* **36,** 1813-1825

Shrader, T.E., Tobias, J.W., and Varshavsky, A. (1993) The N-end rule in *Escherichia coli*: cloning and analysis of the leucyl, phenylalanyl-tRNA-protein transferase gene *aat. J.Bacteriol.* **175,** 4364-4374

Sorensen, M.A., Elf, J., Bouakaz, E., Tenson, T., Sanyal, S., Bjork, G.R., and Ehrenberg, M. (2005) Over expression of a tRNA(Leu) isoacceptor changes charging pattern of leucine tRNAs and reveals new codon reading. *J.Mol.Biol.* **354,** 16-24

Stanzel, M., Schon, A., and Sprinzl, M. (1994) Discrimination against misacylated tRNA by chloroplast elongation factor Tu. *Eur.J.Biochem.* **219,** 435-439

Stortchevoi, A., Varshney, U., and RajBhandary, U.L. (2003) Common location of determinants in initiator transfer RNAs for initiator-elongator discrimination in bacteria and in eukaryotes. *J.Biol.Chem.* **278,** 17672-17679

Suto, K., Shimizu, Y., Watanabe, K., Ueda, T., Fukai, S., Nureki, O., and Tomita, K. (2006) Crystal structures of leucyl/phenylalanyl-tRNA-protein transferase and its complex with an aminoacyl-tRNA analog. *EMBO J.* **25,** 5942-5950

Tobias, J.W., Shrader, T.E., Rocap, G., and Varshavsky, A. (1991) The N-end rule in bacteria. *Science.* **254,** 1374-1377

Villet, R., Fonvielle, M., Busca, P., Chemama, M., Maillard, A.P., Hugonnet, J.E., Dubost, L., Marie, A., Josseaume, N., Mesnage, S., Mayer, C., Valery, J.M., Etheve-Quelquejeu, M., and Arthur, M. (2007) Idiosyncratic features in tRNAs participating in bacterial cell wall synthesis. *Nucleic Acids Res.* **35,** 6870-6883

Wagner, A.M., Fegley, M.W., Warner, J.B., Grindley, C.L., Marotta, N.P., and Petersson, E.J. (2011) N-terminal protein modification using simple aminoacyl transferase substrates. *J.Am.Chem.Soc.* **133,** 15139-15147

Wang, K.H., Oakes, E.S., Sauer, R.T., and Baker, T.A. (2008a) Tuning the strength of a bacterial N-end rule degradation signal. *J.Biol.Chem.* **283,** 24600-24607

Wang, K.H., Roman-Hernandez, G., Grant, R.A., Sauer, R.T., and Baker, T.A. (2008b) The molecular basis of N-end rule recognition. *Mol.Cell.* **32,** 406-414

Watanabe, K., Toh, Y., Suto, K., Shimizu, Y., Oka, N., Wada, T., and Tomita, K. (2007) Protein-based peptide-bond formation by aminoacyl-tRNA protein transferase. *Nature.* **449,** 867-871

Wendrich, T.M., Blaha, G., Wilson, D.N., Marahiel, M.A., and Nierhaus, K.H. (2002) Dissection of the mechanism for the stringent factor RelA. *Mol.Cell.* **10,** 779-788

Wolfson, A.D., and Uhlenbeck, O.C. (2002) Modulation of $tRNA^{Ala}$ identity by inorganic pyrophosphatase. *Proc.Natl.Acad.Sci.U.S.A.* **99,** 5965-5970

Zaborske, J.M., Narasimhan, J., Jiang, L., Wek, S.A., Dittmar, K.A., Freimoser, F., Pan, T., and Wek, R.C. (2009) Genome-wide analysis of tRNA charging and activation of the eIF2 kinase Gcn2p. *J.Biol.Chem.* **284,** 25254-25267

Chapter 6

Conclusions and Discussions

"Logic will get you from A to B. Imagination will take you everywhere."

Albert Einstein

A version of this chapter is submitted to:

Fung AW and Fahlman RP (2014) The Molecular Basis of Post-Translational Addition of Amino Acids in *Escherichia coli* N-end Rule Pathway. Submitted.

6.1. Overview

Targeted intracellular proteolysis is a fundamental biological process that removes misfolded or damaged protein as well as controls levels of regulatory proteins as a rapid adaptation mechanism to cellular stress (Gur *et al.* 2011). A protein that possesses a degradation signal (degron), either intrinsically or post-translationally modified, is targeted for protein degradation by processive enzymes such as the 26S proteasome in eukaryotes or Clp proteases in prokaryotes. An intrinsic degron may become exposed by exo- or endo-peptidase (non-processive) proteolytic cleavage or alternatively by association or dissociation with its interaction partners. Post-translational modifications that target proteins for degradation may include acetylation (Hwang *et al.* 2010, Shemorry *et al.* 2013), phosphorylation (Hwang and Varshavsky 2008), oxidation (Zhang *et al.* 1998, Davydov and Varshavsky 2000, Hu *et al.* 2005), deamidation (Baker and Varshavsky 1995) or addition of amino acids (Leibowitz and Soffer 1969, Soffer *et al.* 1969, Gonda *et al.* 1989, Balzi *et al.* 1990, Tobias *et al.* 1991).

Aminoacyl-tRNA protein transferases catalyze the post-translational addition of amino acids from an aa-tRNA onto the N-terminus of a protein. The modified protein is targeted for protein degradation via the N-end rule pathway, which relates the stability of a protein based on the identity of N-terminal amino acid residue. This class of enzyme is highly conserved with enzymes identified from eubacterial to mammals (Bachmair and

Varshavsky 1989, Gonda *et al.* 1989, Tobias *et al.* 1991). Eukaryotic ATE1 plays key roles in a diversity of physiological functions (see review (Saha and Kashina 2011, Gibbs *et al.* 2014)). However there is lack of molecular details with regards to ATE1 structure and mechanisms. The *E. coli* L/F transferase serves as an ideal model for the study of this class of enzymes because of the available X-ray crystal structures and sensitive functional assays developed (Suto *et al.* 2006, Watanabe *et al.* 2007, Ebhardt *et al.* 2009).

$FemX_{Wv}$ from *Weissella viridescens* belongs to the "Dupli-GNAT" protein superfamily with prokaryotic L/F transferase and eukaryotic ATE1 (Rai *et al.* 2006). $FemX_{Wv}$ catalyzes the interpeptide bond formation where an Ala is transferred from Ala-$tRNA^{Ala}$ to the ε–amino group of Lys_3 of the UDP-MurNAc-pentapeptide (Billot-Klein *et al.* 1997). $FemX_{Wv}$ consists of two GNAT-like domains, similar to ATE1, meanwhile L/F transferase consists of one partial GNAT-like domain (Biarrotte-Sorin *et al.* 2004, Dong *et al.* 2007). $FemX_{Wv}$ catalyzes a similar tRNA-dependent post-translational addition of amino acid (peptide bond formation) as L/F transferase and ATE1. Additionally $FemX_{Wv}$ also has a wealth of biochemical and structural data using substrate analogues, including peptide substrate, bi-substrate RNA-peptide conjugate, and product peptide (Biarrotte-Sorin *et al.* 2004, Chemama *et al.* 2009, Fonvielle *et al.* 2013a, Fonvielle *et al.* 2013b, Mellal *et al.* 2013). Despite differences in substrate specificities, $FemX_{Wv}$ may provide insightful comparisons to the

molecular mechanisms by L/F transferase and ATE1. Here we will compare and contrast the molecular basis on the catalytic mechanism, substrate analogue design, and substrate specificities between these enzymes.

6.2. An Alternative Proton Shuttling Catalytic Mechanism

6.2.1. Ribosomal Peptide Bond Formation

The tRNA-dependent post-translational addition of amino acids reaction is analogous to the ribosomal peptide bond formation. The past decade marks significant advances in the study of ribosomal peptide bond formation catalytic mechanism through biochemical, structural and computational approaches (see recent reviews for more details (Leung *et al.* 2011, Pech and Nierhaus 2012)).

Briefly, the elongation step of prokaryotic protein biosynthesis begins with the delivery of an aa-tRNA from the ternary complex with EF-Tu:GTP to the ribosomal A-site. The correct pairing of the cognate codon on the mRNA and anticodon on the tRNA induces GTP hydrolysis, which leads to conformational changes that positions aa-tRNA in the peptidyltransferase center (PTC) in the 50S large subunit of the ribosome while EF-Tu dissociates. Peptide bond formation occurs in the PTC, where the α-amino group of A-site aa-tRNA nucleophilically attacks the carbonyl carbon of the P-site peptidyl-tRNA. The current catalytic mechanism model is described as a double proton shuttling mechanism using the P-

site A_{76} 2'-OH and a water molecule in the active site (Schmeing *et al.* 2005a, Erlacher and Polacek 2008, Schmeing and Ramakrishnan 2009, Wallin and Aqvist 2010, Leung *et al.* 2011). The ribosomal residues do not appear to participate in catalysis directly (Polacek *et al.* 2001, Thompson *et al.* 2001, Beringer *et al.* 2003, Youngman *et al.* 2004, Erlacher *et al.* 2005, Bieling *et al.* 2006, Trobro and Aqvist 2008, Chirkova *et al.* 2010). Instead the ribosome is proposed to play a role in substrate positioning and solvent reorganization through an induced-fit conformational change and an extensive hydrogen bond network (Weinger *et al.* 2004b, Schmeing *et al.* 2005b, Trobro and Aqvist 2005, Trobro and Aqvist 2006, Beringer and Rodnina 2007b). The A_{76} 2'-OH at the 3' end of the P-site aa-tRNA has been proposed to play an important role by acting as a general acid/base in "substrate-assisted" proton shuttling that bridges the α-amino attacking group and 3'-oxygen leaving group through a six- or eight-membered ring transition state (without and with a water molecule) (Weinger *et al.* 2004b, Schmeing *et al.* 2005a, Erlacher *et al.* 2006, Trobro and Aqvist 2006, Weinger and Strobel 2006, Wallin and Aqvist 2010).

However the details with regards to the timing and proton transfer steps of the reaction remains elusive. Specifically, the proton shuttling reaction mechanism may occur via a fully concerted (no intermediates), a two-step (with a T^- intermediate), or a three-step (with $T^{\pm}$ and T^- intermediates, similar to the uncatalyzed reaction) reaction mechanism.

Recent data supports an alternative mechanism suggesting that the ribosome plays a dual role, where not only does the ribosome induced-fit the substrates but also directly participates in the chemistry of peptide bond formation (i.e. a two-step reaction mechanism through a concerted T^- intermediate formation followed by a breakdown to products) (Kingery *et al.* 2008, Wallin and Aqvist 2010, Hiller *et al.* 2011, Kuhlenkoetter *et al.* 2011, Byun and Kang 2013).

6.2.2. Non-ribosomal Peptide Bond Formation by L/F transferase

In the case for L/F transferase, an initial protein-based catalytic mechanism has been proposed based on available complex structures with substrate analogues and product peptide (Suto *et al.* 2006, Watanabe *et al.* 2007). However an alternative mechanism is proposed based on additional mutagenesis and more sensitive data-collection (**Chapter 3**). L/F transferase, like the ribosome, does not participate in the peptide bond formation chemistry directly but catalyzes the reaction by binding and orientating the substrates (Fung *et al.* 2011). This alternative mechanism mirrors the proton-shuttling mechanism that has been described for the ribosomes (Weinger and Strobel 2006, Beringer and Rodnina 2007b, Fung *et al.* 2011). The alternative catalytic mechanism illustrated in **Figure 6-1** involves the participation of an aa-tRNA A_{76} 2'-OH where it contributes to catalysis by acting as a general acid/base in proton shuttling, while L/F transferase has a more passive role in the specific binding and positioning of the substrates (Fung *et al.* 2011).

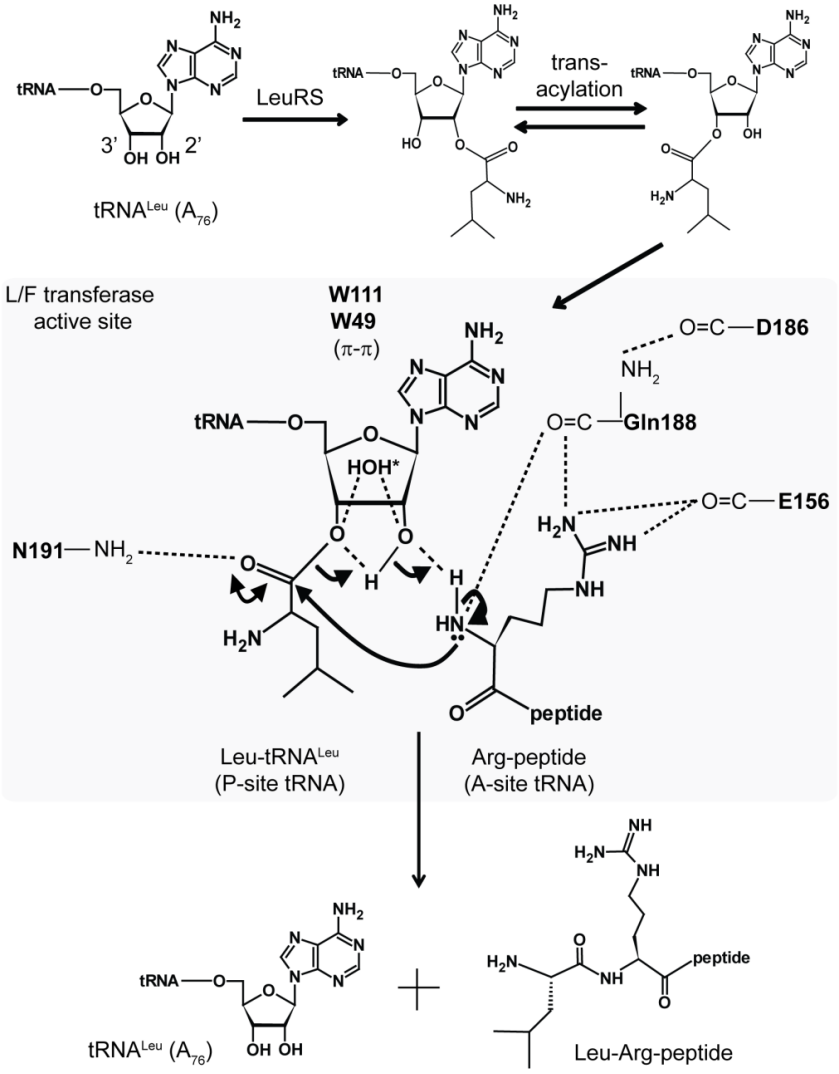

Figure 6-1: A proposed catalytic mechanism of tRNA-dependent peptide bond formation catalyzed by L/F transferase. The $tRNA^{Leu}$ substrate is first aminoacylated by LeuRS at the 2' position, which undergoes simultaneous transacylation to the 3' position that binds into the active site of L/F transferase. Active site residues form hydrogen bonds (E156, D186, Q188, N191) and π-π interactions (W49 and W111) to the donor Leu-$tRNA^{Leu}$ and acceptor Arg-peptide substrates for optimal binding and positioning. The catalytic mechanism is proposed to be a substrate-assisted proton shuttling mechanism via the formation of a six- or eight-membered (with the asterisk water) ring transition state with the donor 2'-OH, similar to the one proposed for the ribosome **(Schmeing *et al.* 2005a)**. The products of the reaction, deacylated $tRNA^{Leu}$ and Leu-Arg-peptide, are released.

Specifically, *E. coli* LeuRS (class I) and PheRS (unique class II) are known to aminoacylate the 2'-OH of A_{76} of their respective tRNA (Eriani *et al.* 1990). Only the 3'-Phe-$tRNA^{Phe}$ (with a 2'-deoxy group locking the amino acid at the 3' position) can serve as an amino acid donor substrate, meanwhile the 2'-Phe-$tRNA^{Phe}$ (with a 3'-deoxy group locking the amino acid at the 2' position) is not a substrate for L/F transferase *in vitro* (Watanabe *et al.* 2007). Thus similar to the ribosome, the 3'-aa-tRNA isomer is the active substrate for L/F transferase. This implies that trans-acylation of aa-tRNA is required, either prior entrance into or within the active site of L/F transferase. This is supported by the observation that the adenosine of rA-Phe is relatively mobile in the crystal structure such that rA-Phe exists as a 1:1 mixture of 2' and 3' isomers (Watanabe *et al.* 2007). The donor aa-tRNA is analogous to the P-site tRNA and the acceptor Arg-peptide substrate is analogous to the A-site tRNA. E156, D186, Q188, and N191 form a hydrogen bond network that positions the two substrates in an optimal orientation (Watanabe *et al.* 2007, Fung *et al.* 2011). The adenine base is stabilized by π-π stacking interaction with W49, which is further stabilized by W111 (Suto *et al.* 2006, Watanabe *et al.* 2007). The "substrate-assisted" nucleophilic attack of α-amino group of N-terminal Arg-peptide onto the carbonyl carbon of aa-tRNA is proposed to occur through a six-membered ring transition state proton-shuttling mechanism where the aa-tRNA A_{76} 2'-OH act as a general acid/base. We

do not rule out the possibility that it may form an eight-membered ring transition state with an active site water molecule.

However some details of the L/F transferase catalytic mechanism remain elusive. Enzyme kinetics involving two substrates and two products (bi-bi) can further be characterized into ping-pong bi-bi, ordered sequential bi-bi, or random order bi-bi mechanism. Ping-pong bi-bi mechanism involves a binary complex, where the enzyme reacts with one substrate to form a product and a modified enzyme, and subsequently reacts with the second substrate to form a second product and regenerats the enzyme. Both ordered sequential bi-bi and random order bi-bi forms a ternary complex with the enzyme, with the ordered mechanism requiring a compulsory order of substrate binding to the enzyme while the random order mechanism may bind to either substrate first. Primary or Lineweaver-Burk plots (plotting reciprocal of initial rate of product formation against reciprocal of substrate 1 concentration while varies substrate 2 concentration) can be used to determine the mechanism between ping-pong (parallel lines that resembles uncompetitive inhibition) and ordered/random ordered (intersecting lines that resembles mixed inhibition). Additionally, the product inhibition patterns observed (i.e. measuring varying concentrations of products A and BX inhibition mode against concentrations of substrates AX and B) can further determine the whether it is an ordered or random order mechanism for the enzyme. Thus, whether the binding of the substrates to L/F transferase occurs via

ping-pong bi-bi, ordered sequential bi-bi or random order bi-bi mechanism remains to be explored.

It has been noted that binding and crystallization experiments of L/F transferase with the substrate peptide alone have not been successful thus far (Watanabe *et al.* 2007, Fung *et al.* 2011), meanwhile tRNA substrate analogues binding and crystallization have been documented (Suto *et al.* 2006, Watanabe *et al.* 2007). The product peptide bound structure also show an absence of the deacylated tRNA product (Watanabe *et al.* 2007). Together this suggests a requisite order of an aa-tRNA binding to L/F transferase first followed by the substrate peptide, and the release of the deacylated tRNA followed by the product peptide. Systematic product inhibition kinetic analyses may decipher between the alternative binding mechanisms.

Additionally, whether the catalytic mechanism by L/F transferase occurs via a fully concerted mechanism or through stable intermediates requires further investigations. More sensitive assays studying the pre-steady state kinetic rate constants and deuterium kinetic isotope effects on the binding, catalysis, and bond-forming/breaking processes may provide more insights into the function of the 2'-OH in this non-ribosomal peptide bond formation (addition of amino acid) catalytic mechanism.

6.2.3. Non-ribosomal Peptide Bond Formation by FemX$_{Wv}$

A similar protein-based general acid/base catalytic mechanism has also initially been proposed for $FemX_{Wv}$, where D109 and E320 are proposed to be the general acid/base involved in catalysis (Hegde and Shrader 2001, Hegde and Blanchard 2003). The general base D109 is proposed to deprotonate α-NH_3^+ prior to nucleophilic attack and the general acid E320 protonates the 3'-leaving group in a similar fashion as the reverse acylation reaction. However, the $FemX_{Wv}$ structure with UDP-MurNAc-pentapeptide shows that D109 cannot be a catalytic residue as it is not near the active site (Biarrotte-Sorin *et al.* 2004).

Product inhibition kinetic analyses suggest that it is an ordered bi-bi mechanism with a sequential binding of the substrates UDP-MurNAc-pentapeptide and Ala-$tRNA^{Ala}$ to $FemX_{Wv}$, followed by a sequential release of the products $tRNA^{Ala}$ and UDP-MurNAc-hexapeptide (Hegde and Blanchard 2003). tRNA substrate analogues with adjacent deoxyadenosine and non-hydrolyzable aminoacyl amide, triazole and oxadiazole analogues were used to understand the regiospecificity of 2' or 3' aa-tRNA for $FemX_{Wv}$ (Chemama *et al.* 2009, Fonvielle *et al.* 2009, Fonvielle *et al.* 2010, Mellal *et al.* 2013). It has been shown that while $FemX_{Wv}$ binds to both isomers, it only utilizes the 2'-isomer as substrate (Fonvielle *et al.* 2010). This is in contrast to the 3'-isomer utilized by both the ribosome and L/F transferase. Since AlaRS aminoacylates at the 3' position and $FemX_{Wv}$ utilizes the 2'-Ala-$tRNA^{Ala}$ implies that transacylation occur prior to entrance into or within the active site (Fonvielle *et al.* 2010).

Also the removal of the 3'-OH group affects the turnover rate but not substrate binding, suggesting the 3'-OH group plays an important role in a similar proton shuttling catalytic mechanism (Fonvielle *et al.* 2010).

Indeed, the recently available complex structures with bi-substrate peptidyl-RNA conjugated analogue and product Mur-NAc-hexapeptide allow an alternative proton shuttling catalytic mechanism to be proposed for $FemX_{Wv}$ (Fonvielle *et al.* 2013a, Fonvielle *et al.* 2013b). The proposed mechanism describes a lack of overall protein conformational change between the acceptor, bi-substrate, and product complex structures except the β5-β6 loops that flip into the catalytic activity upon binding to bi-substrates (Fonvielle *et al.* 2013a). The proposed catalytic mechanism of $FemX_{Wv}$ is a combination of protein-induced tetrahedral intermediate stabilization and "substrate-assisted" proton shuttling mechanism using the A_{76} 3'-OH of aa-tRNA as the general acid/base (Fonvielle *et al.* 2013a).

Together, the proposed catalytic mechanisms for peptide bond formation by the ribosome and non-ribosomal L/F transferase and FemXWv points to a proton-shuttling mechanism using the vicinal hydroxyls of A_{76} of an aa-tRNA. The 2'-OH group (in ribosome and L/F transferase) and 3'-OH group (in $FemX_{Wv}$) of an aa-tRNA participates in the catalytic mechanism by acting as a general acid/base via a six- or eight-membered ring transition state. However, specific details regarding the timing, number of intermediates and proton transfer steps for all three

mechanisms remain elusive. Additionally L/F transferase and $FemX_{Wv}$ not only share a similar GNAT-like domain, but also share significant similarities in the proton shuttling mechanism. The tRNA-dependent non-ribosomal peptide bond formation catalytic mechanisms by L/F transferase and $FemX_{Wv}$ could be more similar to the mechanism proposed for the ribosome than previously believed.

6.3. The Design of an Improved Substrate Analogue

6.3.1. Chemical Modifications for aa-tRNA Substrate Analogues

The 3' terminal adenosine (A_{76}) of an aa-tRNA has a unique feature where two unmodified hydroxyl groups exist adjacent to each other. The constraining ribose ring, electron-withdrawing ribose oxygen, and vicinal hydroxyl groups increase the chemical reactivity of these hydroxyl groups (pK_as reduces from ~16 (of ethanol) to ~12.5) (Izatt *et al.* 1966, Velikyan *et al.* 2001). The aminoacyl residue is linked via an ester linkage to the 2' or 3'-OH of A_{76} of natural aa-tRNAs and each isomer is readily converted to the other via trans-acylation (also trans-esterification). It has been demonstrated that L/F transferase only utilize 3'-aa-tRNA (Watanabe *et al.* 2007) and the adjacent 2'-OH has been suggested to be involved in a 'substrate-assisted' proton shuttling catalytic mechanism similar to the mechanism described for the ribosome (Fung *et al.* 2011).

Inert aa-tRNA substrate analogues can be generated by chemical modifications via two different approaches. One approach is to substitute

the adjacent 2'-OH group on A_{76} with hydrogen (dA_{76}), a fluoro- (fA_{76}) or a methoxy-group creating a non-isomerizable isomer that maintains the labile ester bond linkage but removes the adjacent hydroxyl. This however may modify the ribose conformation (Weinger *et al.* 2004a, Weinger *et al.* 2004b, Schmeing *et al.* 2005a, Koch *et al.* 2008). The other approach is to modify the chemically active ester linkage connecting the aminoacyl moiety to a stable bond, while retaining the adjacent hydroxyl group. Substitutions such as amide, phosphate, phosphoramidate, oxadiazole ring, and triazole ring have been used as aa-tRNA substrate analogues.

The amide linkage is non-hydrolyzable and is best exemplified by puromycin and analogues that have been used in studying the stereo- and regio-chemical properties of ribosomal and non-ribosomal peptide bond formation (Welch *et al.* 1995, Schmeing *et al.* 2005a, Suto *et al.* 2006, Beringer and Rodnina 2007a, Charafeddine *et al.* 2007, Zhong and Strobel 2008, Mellal *et al.* 2013). The phosphate and phosphoramidate groups are often used to study the tetrahedral transition state formed during catalysis (Nissen *et al.* 2000, Schmeing *et al.* 2005a, Schmeing *et al.* 2005b, Cressina *et al.* 2009). Meanwhile the oxadiazole and triazole ring substitutions are stable isostere of ester, where they maintain the geometries and stereoelectronic properties of ester linkage but add an additional carbon (Chemama *et al.* 2007, Chemama *et al.* 2009, Fonvielle *et al.* 2010, Fonvielle *et al.* 2013b).

6.3.2. Probing Substrate Analogue Binding to L/F transferase

Puromycin and analogues with an amide linkage were used to probe the molecular basis of tRNA binding to L/F transferase (**Chapter 4**) (Fung *et al.* 2014a). It has been demonstrated that L/F transferase bind to rA-Phe-amide (K_i = 659 μM, minimal substrate analogue rA-Phe with an amide linkage) and puromycin (K_i = 425 μM) with the same order of affinity due to a combined effect of positive binding interactions (increase hydrophobic contacts) with the dimethyl-modified adenine base and negative binding interactions (steric hindrance with M144) with the methoxy-modified phenylalanyl moiety of puromycin (Fung *et al.* 2014a). A comparison of the determined apparent K_i of puromycin (425 μM) and rA-Phe-amide (659 μM) from this study (Fung *et al.* 2014a) and the reported apparent K_M of rA-Phe (124 μM) (Wagner *et al.* 2011) suggests that the amide bond rigidity of puromycin and rA-Phe-amide significantly disfavour its binding to L/F transferase. This is also supported by the observation that a mobile adenosine may be important for binding and catalysis (Watanabe *et al.* 2007).

A non-competitive mode of inhibition has been reported by tRNA substrate analogues (puromycin and rA-Phe-amide) while measuring against acceptor peptide substrate concentrations suggesting that the a_1 and d_1 pockets do not overlap significantly (Fung *et al.* 2014a). Due to the complex nature of two-substrate reactions, a particular inhibitor that is theoretically competitive in nature may not necessarily result in a

characteristic competitive inhibition pattern (Palmer 1991). These findings provide insight for the design of substrate analogues with enhanced binding for studying the 3' end of aa-tRNA recognition.

6.3.3. A Proposed Improved Substrate Analogue for L/F transferase

Recently $FemX_{Wv}$ alanyl transferase has been shown to be efficiently inhibited by a peptidyl-RNA conjugate analogue to picomolar concentrations and the analogue-bound complex crystal structure has been solved (Fonvielle *et al.* 2013a, Fonvielle *et al.* 2013b). The peptidyl-RNA conjugate analogue is a result of the peptide substrate and aa-tRNA substrate linked together by a 1, 4-triazole ring (Fonvielle *et al.* 2013b). This strategy in utilizing both substrates significantly increases its binding affinity and inhibitory effects to the $FemX_{Wv}$ enzyme (comparing *p*M of the bisubstrate analogue to low μM of amide, triazole and oxadiazole ring analogues) (Chemama *et al.* 2007, Fonvielle *et al.* 2010, Fonvielle *et al.* 2013b, Mellal *et al.* 2013). A similar strategy in targeted inhibitor design may be useful for L/F transferase and other tRNA-dependent "Dupli-GNAT" superfamily proteins.

Here we propose a possible improved substrate analogue for L/F transferase based on the molecular insights from this thesis (**Figure 6-2**). This improved substrate analogue would include the following. **1)** A donor aa-tRNA substrate component with a $tRNA^{Leu}$ (5'-CAG-3') acceptor stem mini helix (24 nucleotide) (**Chapter 5**), and dimethyl modification on the adenine base of A_{76} (**Chapter 4**). **2)** An acceptor peptide substrate

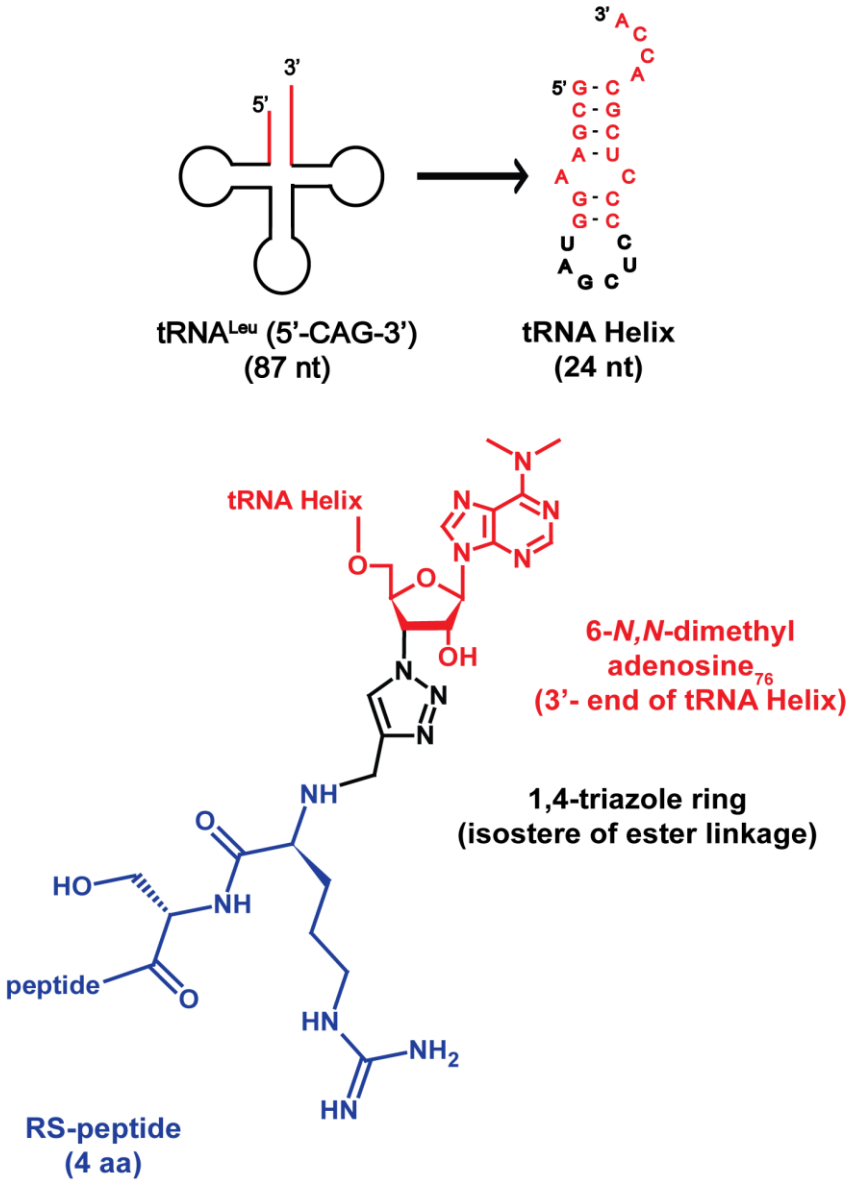

Figure 6-2: Chemical structure of a potential substrate analogue designed for L/F transferase. The designed substrate analogue combines the optimal substrate specificity for L/F transferase including 1) acceptor stem tRNA helix of $tRNA^{Leu}$ (CAG), 2) dimethyl modification on the 3'-A_{76}, 3) acceptor peptide with at least four amino acids in length with an N-terminal Arg (a_1) and a penultimate Ser (a_2) residue, and 4) the two substrates are linked by a 1,4-triazole ring that is an isostere of ester linkage. A Leu residue (d_1) may be further substituted on to the triazole ring for more affinity.

component with at least 4 amino acids in length, an N-terminal Arg for the a_1 pocket, and a favourable penultimate residue at the a_2 pocket (i.e. Ser). **3)** And the two substrate components are covalently linked through a 1, 4-triazole ring (or other isostere of ester linkage). A leucyl group may further be substituted to the triazole ring to mimic Leu in the d_1 pocket for additional binding affinity. Indeed, Santarem *et al*. have shown to successfully synthesize stable Phe-tRNA^{Phe} and Leu-tRNA^{Leu} analogues (and conjugate to a RNA acceptor stem mini-helix or a full tRNA body via a triazole ring (Santarem *et al.* 2014). This proposed bi-substrate analogue would utilize the affinity of both substrates and enable higher affinity binding. Future investigations may reveal additional chemical modifications to improve the design of substrate analogues for L/F-transferase. An improved substrate analogue may be used for complex crystallization experiments to further understand the substrate binding and specificities for L/F transferase.

6.4. A Proposed tRNA Recognition Model

6.4.1. Recognition of the Acceptor Stem of an aa-tRNA

An atypical role of tRNA is targeted proteolysis, where L/F transferase utilizes aa-tRNAs as an activated source of amino acids for tagging proteins for degradation. Among the possible five leucyl-tRNA isoacceptors in *E. coli*, L/F transferase exhibits a strong preference for the abundant isoacceptor - tRNA^{Leu} (CAG) (Rao and Kaji 1974). *In vitro* transcribed tRNA and tRNA hybrids were used to investigate this

isoacceptor preference (**Chapter 5**). While mutations to the D- and T-stem of the aa-tRNA isoacceptors did not result in altered L/F transferase utilization, mutations to the terminal five base pairs in the acceptor stem alter the rate of reaction (Fung *et al.* 2014b). Through detailed analysis of tRNA hybrids, two independent sequence elements in the acceptor stem of Leu-tRNA^{Leu} (CAG) were identified to be important for optimal binding by L/F transferase - the G_3:C_{70} base pair and a set of four combined nucleotides (C_{68}, A_4:U_{69}, and C_{72}) (Fung *et al.* 2014b). The specific molecular mechanism of the sequence-dependent recognition of the acceptor stem and utilization of the isoacceptor tRNA^{Leu} (CAG) remains to be explored.

Compiling the existing structural and biochemical data, the current tRNA recognition model by L/F transferase includes the recognition of 3'-aminoacyl adenosine through the C-shaped hydrophobic pocket binding the amino acid side chain and π-π stacking with the nucleotide base, two sets of sequence-dependent recognition elements in the acceptor stem of Leu-tRNA interacting with the positive cluster on L/F transferase, and the sequence-independent recognition of the D-stem (Suto *et al.* 2006, Watanabe *et al.* 2007, Fung *et al.* 2014b).

6.4.2. aa-tRNA: A Shared Substrate with the Translation Machinery

Post-translational addition of amino acids using aa-tRNA substrates implies that these tRNA-dependent transferases share common

aminoacyl-tRNA synthetases (aaRSs) and tRNAs with the translational machinery and therefore are predicted to compete with elongation factor Tu (EF-Tu) for aa-tRNAs before their utilization by the ribosome. **Figure 6-3** summarizes the recognition nucleotides of Leu-tRNA^{Leu} (CAG) by *E. coli* L/F transferase, LeuRS (Asahara *et al.* 1993a, Asahara *et al.* 1993b, Asahara *et al.* 1998, Larkin *et al.* 2002), and EF-Tu (Schrader *et al.* 2009, Schrader and Uhlenbeck 2011, Schrader *et al.* 2011). All three enzymes interact with the single stranded 3' CCA end but their recognition nucleotides are independent and distinct from each other, perhaps adding to the evolutionary selective pressures on these sequences. EF-Tu has been demonstrated to have strong (K_D values in the low nM range) yet equivalent affinities to all elongator aa-tRNAs using compensatory roles of the esterified amino acid and three adjacent base pairs (49:65, 50:64, and 51:63) in the T-stem of the tRNA body (Andersen and Wiborg 1994, LaRiviere *et al.* 2001, Schrader *et al.* 2011). Given that L/F transferase have a low *in vivo* concentration (L/F ~0.5 μM versus EF-Tu ~100 μM) with a weak affinity for aa-tRNA (L/F K_D ~200 nM versus EF-Tu K_D ~5 nM) when compared to EF-Tu, this predicts a competition of aa-tRNA substrates that is not in favor for L/F transferase (Leibowitz and Soffer 1969, Scarpulla *et al.* 1976, Shrader *et al.* 1993).

Some mechanisms for specialized aa-tRNAs evading the translation machinery have been described that include channeling aa-tRNA substrates through protein complex with aaRS (i.e. VlmA and VlmL in

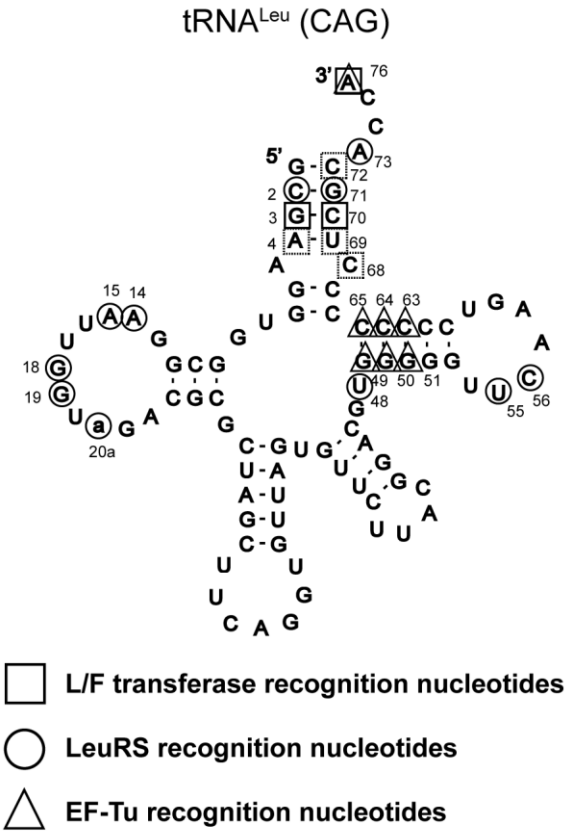

Figure 6-3: A summary of *E. coli* Leu-tRNA^{Leu} (CAG) recognition nucleotides by L/F transferase (□), LeuRS (○), and EF-Tu (Δ). Solid squares represent the major determinant G_3:C_{70} base pair, meanwhile dashed line squares represent the set of four nucleotides (C_{72}, A_4:U_{69}, C_{68}) for L/F transferase aa-tRNA recognition. Data for LeuRS recognition was from (Asahara *et al.* 1993a, Asahara *et al.* 1993b, Asahara *et al.* 1998, Larkin *et al.* 2002), and data for EF-Tu recognition was from (Schrader *et al.* 2009, Schrader and Uhlenbeck 2011, Schrader *et al.* 2011). Nucleotides numbering is according to (Sprinzl *et al.* 1998).

Streptomyces viridifaciens) (Garg *et al.* 2008), and utilizing idiosyncratic features of specific tRNA isoacceptors (i.e. FemXAB in *Staphylococcus aureus*) (Giannouli *et al.* 2009). However there is no data indicating the presence of a specialized LeuRS or $tRNA^{Leu}$ isoacceptor for L/F transferase. It remains somewhat unclear how free canonical aa-tRNA species participate in both translation and alternative functions. One possible model has been proposed; where MprF transferase of *Clostridium perfringens* and EF-Tu have similar affinities for cognate aa-tRNAs suggesting that aa-tRNAs may simultaneous supply amino acids between protein synthesis and alternative functions (Roy and Ibba 2008). Another possible model is through the distinct recognition of aa-tRNA by an anti-parallel coiled-coil domain. This coiled-coil domain (α9-α10) (**Figure 6-4**) has been identified in the glycyl transferase FemA of *Staphylococcus aureus* and provides a platform for tRNA-protein interactions (Biou *et al.* 1994, Benson *et al.* 2002). Similarly, the alanyl/seryl transferase MurM of *Streptococcus pneumonia* has been demonstrated to use this coiled-coil domain to differentiate $tRNA^{Ala}$ and $tRNA^{Ser}$ (Filipe *et al.* 2001, Fiser *et al.* 2003).

6.4.3. Comparing aa-tRNA Recognition Mechanism with $FemX_{Wv}$

However, L/F transferase and $FemX_{Wv}$ do not have this coiled-coil domain and the tRNA-protein interactions are restricted to the globular GNAT-like domain (Biarrotte-Sorin *et al.* 2004, Dong *et al.* 2007). The Ala-$tRNA^{Ala}$ specificity for $FemX_{Wv}$ is mainly through steric hindrance of

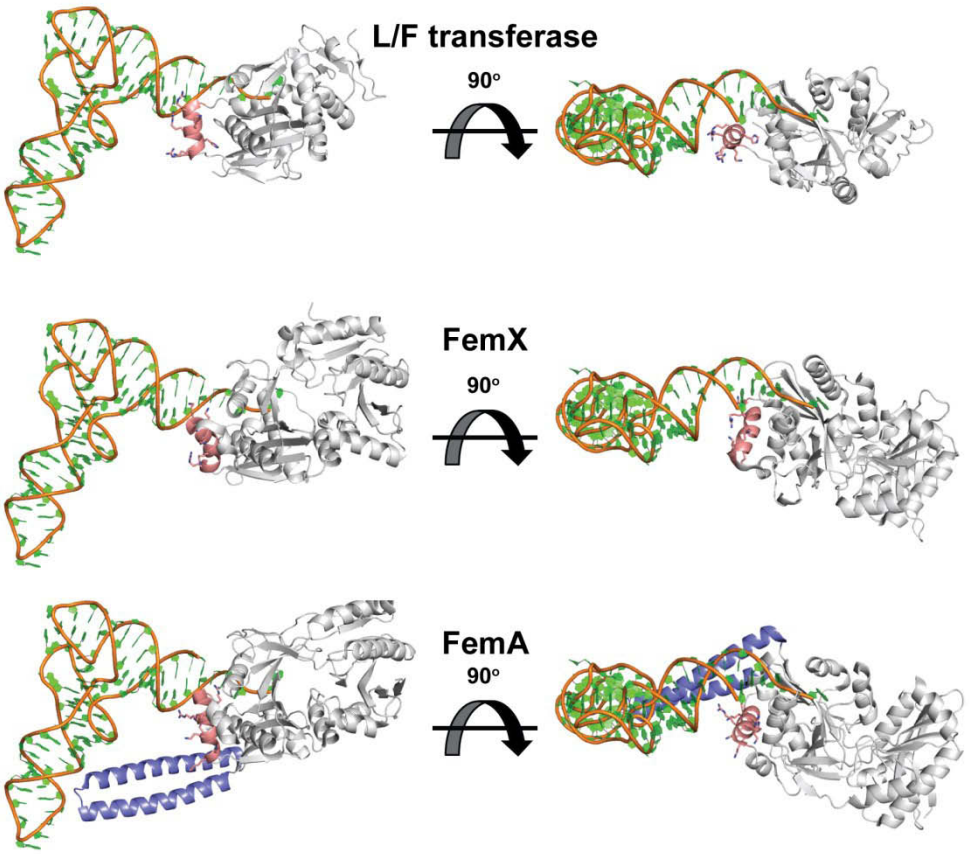

Figure 6-4: A proposed model of aminoacyl-tRNA recognition by GNAT-like domain containing aminoacyl-tRNA protein transferases. The 3' CCA end of the peptidyl-RNA conjugate (PDB ID: 4II9) was used as a reference to dock the yeast $tRNA^{Phe}$ (shown as cartoon, PDB ID: 1EHZ) onto *E. coli L*/F transferase (top panel, cartoon, PDB ID: 2Z3K), *W. viridescens* $FemX_{Wv}$ (middle panel, cartoon, PDB ID: 4II9), and *S. aureus* $FemA_{Sa}$ (bottom panel, cartoon, PDB ID: 1LRZ). Two views of the docking model are shown. The positive residues (shown as sticks) on α2 of L/F transferase, α6 of FemX and α6 of FemA (highlighted in pink) interact with the acceptor stem (terminal five base pairs and terminal two base pairs respectively) of the aminoacyl-tRNA substrate. It has been proposed that the additional coiled-coil region (α9-10, highlighted in blue) of $FemA_{Sa}$ contributes to aminoacyl-tRNA recognition. This RNA-protein interaction may play a role in determining aminoacyl-tRNA specificity of these enzymes. The model was generated using PyMOL version 1.41.

the aminoacyl moiety, where it excludes most amino acids besides glycine (Fonvielle *et al.* 2009). To distinguish $tRNA^{Ala}$ from $tRNA^{Gly}$, $FemX_{Wv}$ recognizes the 2:71 base pair in the acceptor stem specifically ($tRNA^{Ala}$ with the determinant G2:C71 base pair, and $tRNA^{Gly}$ with the anti-determinant C2:G71 base pair) (Villet *et al.* 2007, Fonvielle *et al.* 2009). Thus the tRNA recognition by $FemX_{wv}$ was demonstrated to depend on the terminal two base pairs of the acceptor stem and the main determinant is the G_2:C_{71} base pair (Villet *et al.* 2007, Fonvielle *et al.* 2009). Additionally the identity element for AlaRS (G_3:U_{70}) is not important for recognition, such that $FemX_{wv}$ display distinct recognition sites in the aa-tRNA from AlaRS (Villet *et al.* 2007).

The aa-tRNA recognition by the structurally similar L/F transferase and $FemX_{Wv}$ are more similar than once thought. First L/F transferase also uses steric hindrance by the C-shaped hydrophobic pocket to select for the aminoacyl moiety (Suto *et al.* 2006, Watanabe *et al.* 2007). From this study, we demonstrated that L/F transferase similarly has a major determinant base pair (G_3:C_{70}) for efficient utilization of specific $tRNA^{Leu}$ isoacceptors. Additionally, L/F transferase and $FemX_{Wv}$ recognition in the acceptor stem are distinct from LeuRS or AlaRS recognition, respectively (Asahara *et al.* 1993a, Asahara *et al.* 1993b, Asahara *et al.* 1998, Larkin *et al.* 2002, Fung *et al.* 2014b). However there are difference in the number of contacts between the protein and the aa-tRNA substrate between L/F transferase and $FemX_{Wv}$. $FemX_{Wv}$ is suggested to recognize the distal

end of an aa-tRNA (up to the first two base pairs) (Villet *et al.* 2007, Fonvielle *et al.* 2009), meanwhile our data suggests that L/F transferase recognizes up to five base pairs in the acceptor stem specifically.

6.4.4. A Proposed Model for tRNA Recognition by GNAT-like domain and Evasion of Translation Machinery

Here we propose a tRNA recognition model for GNAT-like domain containing enzymes based on the biochemical and structural data from L/F transferase and $FemX_{Wv}$ and FemA (**Figure 6-4**) (Biou *et al.* 1994, Abramochkin and Shrader 1996, Benson *et al.* 2002, Suto *et al.* 2006, Villet *et al.* 2007, Watanabe *et al.* 2007, Fonvielle *et al.* 2009, Fonvielle *et al.* 2013a, Fung *et al.* 2014b). Modeling of an yeast $tRNA^{Phe}$ (PDB ID: 1EHZ) docked onto L/F transferase (PDB ID: 2Z3K) and $FemX_{Wv}$ (PDB ID: 4II9) shows that the positive amino acid cluster on α2 helix of L/F transferase (pink) and α6 helix of $FemX_{Wv}$ and FemA (pink) in the C-terminal GNAT-like domain interacts with the terminal end of the acceptor stem (5 base pairs for L/F transferase and 2 base pairs for $FemX_{Wv}$) and this RNA-protein interaction may contribute to the acceptor stem specificity observed. For FemA, the presence of a coiled-coil domain (α9-α10) provides additional RNA-protein interactions.

Although this proposal of a basic α helix recognition mechanism may explain the tRNA acceptor stem recognition, however they do not affect the strong binding of aa-tRNA to EF-Tu. L/F transferase and $FemX_{Wv}$ do

not seem to follow any of the above potential evasion mechanisms. Perhaps there are still unidentified factors that aid in aa-tRNA substrate recruitment or these transferases access aa-tRNAs when EF-Tu is inactivated. We hypothesize a general mechanism for L/F transferase

accessing free canonical aa-tRNA (Fung *et al.* 2014b). When amino acids are limited during the stringent response, half of the GTP molecules are converted to pentaphosphate guanosine (pppGpp) (Fiil *et al.* 1972). The loss of GTP molecules would affect the ability of EF-Tu to bind to aa-tRNAs efficiently as it requires a GTP molecule to form a ternary complex. During the stringent response, the pppGpp molecules are further hydrolyzed into tetraphosphate guanosine (ppGpp) by guanosine pentaphosphate phosphatases, the functional molecule of the stringent response (Wu and Xie 2009). The ppGpp molecule resembles GDP and inhibits EF-Tu directly (K_i = 7 x 10^{-7} M) or indirectly via trapping the EF-Tu:EF-Ts cycling complex (K_i = 4 x 10^{-5} M) (Rojas *et al.* 1984). This general EF-Tu inactivation mechanism would allow free aa-tRNA to be used for alternative functions such as the post-translational addition of amino acids.

6.5. Concluding Remarks

Here we investigated the molecular basis of catalytic mechanism (**Chapter 3**), substrate analogue design (**Chapter 4**) and tRNA substrate recognition (**Chapter 5**) by L/F transferase through the use of available X-ray crystal structures, mutagenesis, *in vitro* transcribed tRNAs and tRNA

hybrids, and quantitative matrix-assisted laser desorption/ionization time of flight mass spectrometry functional assay developed by our lab.

In this thesis, we proposed an alternative substrate-assisted proton shuttling catalytic mechanism for L/F transferase that is similar to one proposed for the ribosome. We proposed a potential design of an improved substrate analogue for L/F transferase. And we fine-tuned and proposed a tRNA recognition model by L/F transferase where the positive cluster on $\alpha 2$ helix modulates tRNA recognition through the acceptor stem.

Taken together, our molecular studies into the L/F transferase reaction revolutionize the current understanding of molecular details in the catalytic mechanism, substrate analogue design, and tRNA substrate recognition.

6.6. References

Abramochkin, G., and Shrader, T.E. (1996) Aminoacyl-tRNA recognition by the leucyl/phenylalanyl-tRNA-protein transferase. *J.Biol.Chem.* **271,** 22901-22907

Andersen, C., and Wiborg, O. (1994) *Escherichia coli* elongation-factor-Tu mutants with decreased affinity for aminoacyl-tRNA. *Eur.J.Biochem.* **220,** 739-744

Asahara, H., Himeno, H., Tamura, K., Hasegawa, T., Watanabe, K., and Shimizu, M. (1993a) Recognition nucleotides of *Escherichia coli* $tRNA^{Leu}$ and its elements facilitating discrimination from $tRNA^{Ser}$ and $tRNA^{Tyr}$. *J.Mol.Biol.* **231,** 219-229

Asahara, H., Himeno, H., Tamura, K., Nameki, N., Hasegawa, T., and Shimizu, M. (1993b) Discrimination among *E. coli* tRNAs with a long variable arm. *Nucleic Acids Symp.Ser.* **(29),** 207-208

Asahara, H., Nameki, N., and Hasegawa, T. (1998) *In vitro* selection of RNAs aminoacylated by *Escherichia coli* leucyl-tRNA synthetase. *J.Mol.Biol.* **283,** 605-618

Bachmair, A., and Varshavsky, A. (1989) The degradation signal in a short-lived protein. *Cell.* **56,** 1019-1032

Baker, R.T., and Varshavsky, A. (1995) Yeast N-terminal amidase. A new enzyme and component of the N-end rule pathway. *J.Biol.Chem.* **270,** 12065-12074

Balzi, E., Choder, M., Chen, W.N., Varshavsky, A., and Goffeau, A. (1990) Cloning and functional analysis of the arginyl-tRNA-protein transferase gene ATE1 of *Saccharomyces cerevisiae. J.Biol.Chem.* **265,** 7464-7471

Benson, T.E., Prince, D.B., Mutchler, V.T., Curry, K.A., Ho, A.M., Sarver, R.W., Hagadorn, J.C., Choi, G.H., and Garlick, R.L. (2002) X-ray crystal structure of *Staphylococcus aureus* FemA. *Structure.* **10,** 1107-1115

Beringer, M., Adio, S., Wintermeyer, W., and Rodnina, M. (2003) The G2447A mutation does not affect ionization of a ribosomal group taking part in peptide bond formation. *RNA.* **9,** 919-922

Beringer, M., and Rodnina, M.V. (2007a) Importance of tRNA interactions with 23S rRNA for peptide bond formation on the ribosome: studies with substrate analogs. *Biol.Chem.* **388,** 687-691

Beringer, M., and Rodnina, M.V. (2007b) The ribosomal peptidyl transferase. *Mol.Cell.* **26,** 311-321

Biarrotte-Sorin, S., Maillard, A.P., Delettre, J., Sougakoff, W., Arthur, M., and Mayer, C. (2004) Crystal structures of Weissella viridescens FemX and its complex with UDP-MurNAc-pentapeptide: insights into FemABX family substrates recognition. *Structure.* **12,** 257-267

Bieling, P., Beringer, M., Adio, S., and Rodnina, M.V. (2006) Peptide bond formation does not involve acid-base catalysis by ribosomal residues. *Nat.Struct.Mol.Biol.* **13,** 423-428

Billot-Klein, D., Shlaes, D., Bryant, D., Bell, D., Legrand, R., Gutmann, L., and van Heijenoort, J. (1997) Presence of UDP-N-acetylmuramyl-hexapeptides and -heptapeptides in enterococci and staphylococci after treatment with ramoplanin, tunicamycin, or vancomycin. *J.Bacteriol.* **179,** 4684-4688

Biou, V., Yaremchuk, A., Tukalo, M., and Cusack, S. (1994) The 2.9 A crystal structure of T. thermophilus seryl-tRNA synthetase complexed with tRNA(Ser). *Science.* **263,** 1404-1410

Byun, B.J., and Kang, Y.K. (2013) A mechanistic study supports a two-step mechanism for peptide bond formation on the ribosome. *Phys.Chem.Chem.Phys.* **15,** 14931-14935

Charafeddine, A., Dayoub, W., Chapuis, H., and Strazewski, P. (2007) First synthesis of 2'-deoxyfluoropuromycin analogues: experimental insight into the mechanism of the Staudinger reaction. *Chem. Eur. J.* **13,** 5566-5584

Chemama, M., Fonvielle, M., Arthur, M., Valery, J.M., and Etheve-Quelquejeu, M. (2009) Synthesis of stable aminoacyl-tRNA analogues containing triazole as a bioisoster of esters. *Chemistry.* **15,** 1929-1938

Chemama, M., Fonvielle, M., Villet, R., Arthur, M., Valery, J.M., and Etheve-Quelquejeu, M. (2007) Stable analogues of aminoacyl-tRNA for inhibition of an essential step of bacterial cell-wall synthesis. *J.Am.Chem.Soc.* **129,** 12642-12643

Chirkova, A., Erlacher, M.D., Clementi, N., Zywicki, M., Aigner, M., and Polacek, N. (2010) The role of the universally conserved A2450-C2063 base pair in the ribosomal peptidyl transferase center. *Nucleic Acids Res.* **38,** 4844-4855

Cressina, E., Lloyd, A.J., De Pascale, G., James Mok, B., Caddick, S., Roper, D.I., Dowson, C.G., and Bugg, T.D. (2009) Inhibition of tRNA-dependent ligase MurM from Streptococcus pneumoniae by phosphonate and sulfonamide inhibitors. *Bioorg.Med.Chem.* **17,** 3443-3455

Davydov, I.V., and Varshavsky, A. (2000) RGS4 is arginylated and degraded by the N-end rule pathway *in vitro. J.Biol.Chem.* **275,** 22931-22941

Dong, X., Kato-Murayama, M., Muramatsu, T., Mori, H., Shirouzu, M., Bessho, Y., and Yokoyama, S. (2007) The crystal structure of leucyl/phenylalanyl-tRNA-protein transferase from *Escherichia coli. Protein Sci.* **16,** 528-534

Ebhardt, H.A., Xu, Z., Fung, A.W., and Fahlman, R.P. (2009) Quantification of the post-translational addition of amino acids to proteins by MALDI-TOF mass spectrometry. *Anal.Chem.* **81,** 1937-1943

Eriani, G., Delarue, M., Poch, O., Gangloff, J., and Moras, D. (1990) Partition of tRNA synthetases into two classes based on mutually exclusive sets of sequence motifs. *Nature.* **347,** 203-206

Erlacher, M.D., Lang, K., Shankaran, N., Wotzel, B., Huttenhofer, A., Micura, R., Mankin, A.S., and Polacek, N. (2005) Chemical engineering of the peptidyl transferase center reveals an important role of the 2'-hydroxyl group of A2451. *Nucleic Acids Res.* **33,** 1618-1627

Erlacher, M.D., Lang, K., Wotzel, B., Rieder, R., Micura, R., and Polacek, N. (2006) Efficient ribosomal peptidyl transfer critically relies on the presence of the ribose 2'-OH at A2451 of 23S rRNA. *J.Am.Chem.Soc.* **128,** 4453-4459

Erlacher, M.D., and Polacek, N. (2008) Ribosomal catalysis: the evolution of mechanistic concepts for peptide bond formation and peptidyl-tRNA hydrolysis. *RNA Biol.* **5,** 5-12

Fiil, N.P., von Meyenburg, K., and Friesen, J.D. (1972) Accumulation and turnover of guanosine tetraphosphate in *Escherichia coli. J.Mol.Biol.* **71,** 769-783

Filipe, S.R., Severina, E., and Tomasz, A. (2001) Functional analysis of Streptococcus pneumoniae MurM reveals the region responsible for its specificity in the synthesis of branched cell wall peptides. *J.Biol.Chem.* **276,** 39618-39628

Fiser, A., Filipe, S.R., and Tomasz, A. (2003) Cell wall branches, penicillin resistance and the secrets of the MurM protein. *Trends Microbiol.* **11,** 547-553

Fonvielle, M., Chemama, M., Lecerf, M., Villet, R., Busca, P., Bouhss, A., Etheve-Quelquejeu, M., and Arthur, M. (2010) Decoding the logic of the tRNA regiospecificity of nonribosomal FemX(Wv) aminoacyl transferase. *Angew.Chem.Int.Ed Engl.* **49,** 5115-5119

Fonvielle, M., Chemama, M., Villet, R., Lecerf, M., Bouhss, A., Valery, J.M., Etheve-Quelquejeu, M., and Arthur, M. (2009) Aminoacyl-tRNA recognition by the FemXWv transferase for bacterial cell wall synthesis. *Nucleic Acids Res.* **37,** 1589-1601

Fonvielle, M., Li de La Sierra-Gallay, I., El-Sagheer, A.H., Lecerf, M., Patin, D., Mellal, D., Mayer, C., Blanot, D., Gale, N., Brown, T., van Tilbeurgh, H., Etheve-Quelquejeu, M., and Arthur, M. (2013a) The structure of FemX(Wv) in complex with a peptidyl-RNA conjugate:

mechanism of aminoacyl transfer from Ala-tRNA(Ala) to peptidoglycan precursors. *Angew.Chem.Int.Ed Engl.* **52,** 7278-7281

Fonvielle, M., Mellal, D., Patin, D., Lecerf, M., Blanot, D., Bouhss, A., Santarem, M., Mengin-Lecreulx, D., Sollogoub, M., Arthur, M., and Etheve-Quelquejeu, M. (2013b) Efficient access to peptidyl-RNA conjugates for picomolar inhibition of non-ribosomal FemX(Wv) aminoacyl transferase. *Chemistry.* **19,** 1357-1363

Fung, A.W., Ebhardt, H.A., Abeysundara, H., Moore, J., Xu, Z., and Fahlman, R.P. (2011) An alternative mechanism for the catalysis of peptide bond formation by L/F transferase: substrate binding and orientation. *J.Mol.Biol.* **409,** 617-629

Fung, A.W., Ebhardt, H.A., Krishnakumar, K.S., Moore, J., Xu, Z., Strazewski, P., and Fahlman, R.P. (2014a) Probing the Leucyl/Phenylalanyl tRNA Protein Transferase Active Site with tRNA Substrate Analogues. *Protein Pept.Lett.* **21,** 603-614

Fung, A.W., Leung, C.C., and Fahlman, R.P. (2014b) The determination of $tRNA^{Leu}$ recognition nucleotides for *Escherichia coli* L/F transferase. *RNA.* **20,** 1210-1222

Garg, R.P., Qian, X.L., Alemany, L.B., Moran, S., and Parry, R.J. (2008) Investigations of valanimycin biosynthesis: elucidation of the role of seryl-tRNA. *Proc.Natl.Acad.Sci.U.S.A.* **105,** 6543-6547

Giannouli, S., Kyritsis, A., Malissovas, N., Becker, H.D., and Stathopoulos, C. (2009) On the role of an unusual $tRNA^{Gly}$ isoacceptor in *Staphylococcus aureus. Biochimie.* **91,** 344-351

Gibbs, D.J., Bacardit, J., Bachmair, A., and Holdsworth, M.J. (2014) The eukaryotic N-end rule pathway: conserved mechanisms and diverse functions. *Trends Cell Biol.*

Gonda, D.K., Bachmair, A., Wunning, I., Tobias, J.W., Lane, W.S., and Varshavsky, A. (1989) Universality and structure of the N-end rule. *J.Biol.Chem.* **264,** 16700-16712

Gur, E., Biran, D., and Ron, E.Z. (2011) Regulated proteolysis in Gram-negative bacteria--how and when? *Nat.Rev.Microbiol.* **9,** 839-848

Hegde, S.S., and Blanchard, J.S. (2003) Kinetic and mechanistic characterization of recombinant *Lactobacillus viridescens* FemX (UDP-N-acetylmuramoyl pentapeptide-lysine N_6-alanyltransferase). *J.Biol.Chem.* **278,** 22861-22867

Hegde, S.S., and Shrader, T.E. (2001) FemABX family members are novel nonribosomal peptidyltransferases and important pathogen-specific drug targets. *J.Biol.Chem.* **276,** 6998-7003

Hiller, D.A., Singh, V., Zhong, M., and Strobel, S.A. (2011) A two-step chemical mechanism for ribosome-catalysed peptide bond formation. *Nature.* **476,** 236-239

Hu, R.G., Sheng, J., Qi, X., Xu, Z., Takahashi, T.T., and Varshavsky, A. (2005) The N-end rule pathway as a nitric oxide sensor controlling the levels of multiple regulators. *Nature.* **437,** 981-986

Hwang, C.S., Shemorry, A., and Varshavsky, A. (2010) N-terminal acetylation of cellular proteins creates specific degradation signals. *Science.* **327,** 973-977

Hwang, C.S., and Varshavsky, A. (2008) Regulation of peptide import through phosphorylation of Ubr1, the ubiquitin ligase of the N-end rule pathway. *Proc.Natl.Acad.Sci.U.S.A.* **105,** 19188-19193

Izatt, R.M., Rytting, J.H., Hansen, L.D., and Christensen, J.J. (1966) Thermodynamics of proton dissociation in dilute aqueous solution. V. An entropy titration study of adenosine, pentoses, hexoses, and related compounds. *J.Am.Chem.Soc.* **88,** 2641-2645

Kingery, D.A., Pfund, E., Voorhees, R.M., Okuda, K., Wohlgemuth, I., Kitchen, D.E., Rodnina, M.V., and Strobel, S.A. (2008) An uncharged amine in the transition state of the ribosomal peptidyl transfer reaction. *Chem.Biol.* **15,** 493-500

Koch, M., Huang, Y., and Sprinzl, M. (2008) Peptide-bond synthesis on the ribosome: no free vicinal hydroxy group required on the terminal ribose residue of peptidyl-tRNA. *Angew.Chem.Int.Ed Engl.* **47,** 7242-7245

Kuhlenkoetter, S., Wintermeyer, W., and Rodnina, M.V. (2011) Different substrate-dependent transition states in the active site of the ribosome. *Nature.* **476,** 351-354

LaRiviere, F.J., Wolfson, A.D., and Uhlenbeck, O.C. (2001) Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. *Science.* **294,** 165-168

Larkin, D.C., Williams, A.M., Martinis, S.A., and Fox, G.E. (2002) Identification of essential domains for *Escherichia coli* tRNA(leu) aminoacylation and amino acid editing using minimalist RNA molecules. *Nucleic Acids Res.* **30,** 2103-2113

Leibowitz, M.J., and Soffer, R.L. (1969) A soluble enzyme from *Escherichia coli* which catalyzes the transfer of leucine and phenylalanine from tRNA to acceptor proteins. *Biochem.Biophys.Res.Commun.* **36,** 47-53

Leung, E.K., Suslov, N., Tuttle, N., Sengupta, R., and Piccirilli, J.A. (2011) The mechanism of peptidyl transfer catalysis by the ribosome. *Annu.Rev.Biochem.* **80,** 527-555

Mellal, D., Fonvielle, M., Santarem, M., Chemama, M., Schneider, Y., Iannazzo, L., Braud, E., Arthur, M., and Etheve-Quelquejeu, M. (2013) Synthesis and biological evaluation of non-isomerizable analogues of Ala-tRNA(Ala). *Org.Biomol.Chem.* **11,** 6161-6169

Nissen, P., Hansen, J., Ban, N., Moore, P.B., and Steitz, T.A. (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science.* **289,** 920-930

Palmer, T. (1991) Understanding Enzymes Third Edition. Ellis Horwood Limited.

Pech, M., and Nierhaus, K.H. (2012) The thorny way to the mechanism of ribosomal peptide-bond formation. *Chembiochem.* **13,** 189-192

Polacek, N., Gaynor, M., Yassin, A., and Mankin, A.S. (2001) Ribosomal peptidyl transferase can withstand mutations at the putative catalytic nucleotide. *Nature.* **411,** 498-501

Rai, R., Mushegian, A., Makarova, K., and Kashina, A. (2006) Molecular dissection of arginyltransferases guided by similarity to bacterial peptidoglycan synthases. *EMBO Rep.* **7,** 800-805

Rao, P.M., and Kaji, H. (1974) Utilization of isoaccepting leucyl-tRNA in the soluble incorporation system and protein synthesizing systems from *E.coli. FEBS Lett.* **43,** 199-202

Rojas, A.M., Ehrenberg, M., Andersson, S.G., and Kurland, C.G. (1984) ppGpp inhibition of elongation factors Tu, G and Ts during polypeptide synthesis. *Mol.Gen.Genet.* **197,** 36-45

Roy, H., and Ibba, M. (2008) RNA-dependent lipid remodeling by bacterial multiple peptide resistance factors. *Proc.Natl.Acad.Sci.U.S.A.* **105,** 4667-4672

Saha, S., and Kashina, A. (2011) Posttranslational arginylation as a global biological regulator. *Dev.Biol.* **358,** 1-8

Santarem, M., Fonvielle, M., Sakkas, N., Laisne, G., Chemama, M., Herbeuval, J.P., Braud, E., Arthur, M., and Etheve-Quelquejeu, M. (2014) Synthesis of 3'-triazoyl-dinucleotides as precursors of stable Phe-tRNA(Phe) and Leu-tRNA(Leu) analogues. *Bioorg.Med.Chem.Lett.* **24,** 3231-3233

Scarpulla, R.C., Deutch, C.E., and Soffer, R.L. (1976) Transfer of methionyl residues by leucyl, phenylalanyl-tRNA-protein transferase. *Biochem.Biophys.Res.Commun.* **71,** 584-589

Schmeing, T.M., Huang, K.S., Kitchen, D.E., Strobel, S.A., and Steitz, T.A. (2005a) Structural insights into the roles of water and the 2' hydroxyl of the P site tRNA in the peptidyl transferase reaction. *Mol.Cell.* **20,** 437-448

Schmeing, T.M., Huang, K.S., Strobel, S.A., and Steitz, T.A. (2005b) An induced-fit mechanism to promote peptide bond formation and exclude hydrolysis of peptidyl-tRNA. *Nature.* **438,** 520-524

Schmeing, T.M., and Ramakrishnan, V. (2009) What recent ribosome structures have revealed about the mechanism of translation. *Nature.* **461,** 1234-1242

Schrader, J.M., Chapman, S.J., and Uhlenbeck, O.C. (2009) Understanding the sequence specificity of tRNA binding to elongation factor Tu using tRNA mutagenesis. *J.Mol.Biol.* **386,** 1255-1264

Schrader, J.M., Chapman, S.J., and Uhlenbeck, O.C. (2011) Tuning the affinity of aminoacyl-tRNA to elongation factor Tu for optimal decoding. *Proc.Natl.Acad.Sci.U.S.A.* **108,** 5215-5220

Schrader, J.M., and Uhlenbeck, O.C. (2011) Is the sequence-specific binding of aminoacyl-tRNAs by EF-Tu universal among bacteria?. *Nucleic Acids Res.* **39,** 9746-9758

Shemorry, A., Hwang, C.S., and Varshavsky, A. (2013) Control of protein quality and stoichiometries by N-terminal acetylation and the N-end rule pathway. *Mol.Cell.* **50,** 540-551

Shrader, T.E., Tobias, J.W., and Varshavsky, A. (1993) The N-end rule in *Escherichia coli*: cloning and analysis of the leucyl, phenylalanyl-tRNA-protein transferase gene *aat. J.Bacteriol.* **175,** 4364-4374

Soffer, R.L., Horinishi, H., and Leibowitz, M.J. (1969) The aminoacyl tRNA-protein transferases. *Cold Spring Harb.Symp.Quant.Biol.* **34,** 529-533

Suto, K., Shimizu, Y., Watanabe, K., Ueda, T., Fukai, S., Nureki, O., and Tomita, K. (2006) Crystal structures of leucyl/phenylalanyl-tRNA-protein transferase and its complex with an aminoacyl-tRNA analog. *EMBO J.* **25,** 5942-5950

Thompson, J., Kim, D.F., O'Connor, M., Lieberman, K.R., Bayfield, M.A., Gregory, S.T., Green, R., Noller, H.F., and Dahlberg, A.E. (2001) Analysis of mutations at residues A2451 and G2447 of 23S rRNA in the peptidyltransferase active site of the 50S ribosomal subunit. *Proc.Natl.Acad.Sci.U.S.A.* **98,** 9002-9007

Tobias, J.W., Shrader, T.E., Rocap, G., and Varshavsky, A. (1991) The N-end rule in bacteria. *Science.* **254,** 1374-1377

Trobro, S., and Aqvist, J. (2005) Mechanism of peptide bond synthesis on the ribosome. *Proc.Natl.Acad.Sci.U.S.A.* **102,** 12395-12400

Trobro, S., and Aqvist, J. (2006) Analysis of predictions for the catalytic mechanism of ribosomal peptidyl transfer. *Biochemistry.* **45,** 7049-7056

Trobro, S., and Aqvist, J. (2008) Role of ribosomal protein L27 in peptidyl transfer. *Biochemistry.* **47,** 4898-4906

Velikyan, I., Acharya, S., Trifonova, A., Foldesi, A., and Chattopadhyaya, J. (2001) The pK(a)'s of 2'-hydroxyl group in nucleosides and nucleotides. *J.Am.Chem.Soc.* **123,** 2893-2894

Villet, R., Fonvielle, M., Busca, P., Chemama, M., Maillard, A.P., Hugonnet, J.E., Dubost, L., Marie, A., Josseaume, N., Mesnage, S., Mayer, C., Valery, J.M., Etheve-Quelquejeu, M., and Arthur, M. (2007) Idiosyncratic features in tRNAs participating in bacterial cell wall synthesis. *Nucleic Acids Res.* **35,** 6870-6883

Wagner, A.M., Fegley, M.W., Warner, J.B., Grindley, C.L., Marotta, N.P., and Petersson, E.J. (2011) N-terminal protein modification using simple aminoacyl transferase substrates. *J.Am.Chem.Soc.* **133,** 15139-15147

Wallin, G., and Aqvist, J. (2010) The transition state for peptide bond formation reveals the ribosome as a water trap. *Proc.Natl.Acad.Sci.U.S.A.* **107,** 1888-1893

Watanabe, K., Toh, Y., Suto, K., Shimizu, Y., Oka, N., Wada, T., and Tomita, K. (2007) Protein-based peptide-bond formation by aminoacyl-tRNA protein transferase. *Nature.* **449,** 867-871

Weinger, J.S., Kitchen, D., Scaringe, S.A., Strobel, S.A., and Muth, G.W. (2004a) Solid phase synthesis and binding affinity of peptidyl transferase transition state mimics containing 2'-OH at P-site position A76. *Nucleic Acids Res.* **32,** 1502-1511

Weinger, J.S., Parnell, K.M., Dorner, S., Green, R., and Strobel, S.A. (2004b) Substrate-assisted catalysis of peptide bond formation by the ribosome. *Nat.Struct.Mol.Biol.* **11,** 1101-1106

Weinger, J.S., and Strobel, S.A. (2006) Participation of the tRNA A76 hydroxyl groups throughout translation. *Biochemistry.* **45,** 5939-5948

Welch, M., Chastang, J., and Yarus, M. (1995) An inhibitor of ribosomal peptidyl transferase using transition-state analogy. *Biochemistry.* **34,** 385-390

Wu, J., and Xie, J. (2009) Magic spot: (p) ppGpp. *J.Cell.Physiol.* **220,** 297-302

Youngman, E.M., Brunelle, J.L., Kochaniak, A.B., and Green, R. (2004) The active site of the ribosome is composed of two layers of conserved nucleotides with distinct roles in peptide bond formation and peptide release. *Cell.* **117,** 589-599

Zhang, N., Donnelly, R., and Ingoglia, N.A. (1998) Evidence that oxidized proteins are substrates for N-terminal arginylation. *Neurochem.Res.* **23,** 1411-1420

Zhong, M., and Strobel, S.A. (2008) Synthesis of isotopically labeled P-site substrates for the ribosomal peptidyl transferase reaction. *J.Org.Chem.* **73,** 603-611

Appendix 1

The Identification of L/F transferase Interacting Partners by Affinity Purification Coupled in-gel LC-MS/MS

Note: David Kramer (BIOCH 499) contributed to the affinity purification coupled LC-MS/MS protein identification procedure (AP-MS).

A1.1. Introduction

L/F transferase activity and substrate specificity were documented over 40 years ago (Leibowitz and Soffer 1969, Soffer 1973). However, the low *in vivo* concentration of L/F transferase (~0.5 μM) has hindered its characterization (Leibowitz and Soffer 1969, Scarpulla *et al.* 1976, Shrader *et al.* 1993). Since Abramochkin *et al*. has constructed the recombinant N-terminal GST-tagged and hexa-histidine-tagged L/F transferase and demonstrated that they are as active as the purified wild-type enzymes (Abramochkin and Shrader 1995), it greatly enhanced the progress of L/F transferase *in vitro* studies (Suto *et al.* 2006, Watanabe *et al.* 2007, Ebhardt *et al.* 2009, Fung *et al.* 2011, Wagner *et al.* 2011, Kawaguchi *et al.* 2013, Fung *et al.* 2014a, Fung *et al.* 2014b). The choice of N-terminal fusion junction was based on the observation that wild-type L/F transferase allowed modification of this region without significant loss of activity (Shrader *et al.* 1993).

Despite recent advances in *in vitro* studies regarding L/F transferase's structure (Suto *et al.* 2006, Dong *et al.* 2007, Watanabe *et al.* 2007), catalytic mechanisms (**Chapter 3**) (Fung *et al.* 2011), substrate specificities (**Chapter 4** and **5**) (Kawaguchi *et al.* 2013, Fung *et al.* 2014a, Fung *et al.* 2014b), and substrate identification (Ninnis *et al.* 2009, Schmidt *et al.* 2009, Humbard *et al.* 2013), currently there is no validated protein that interacts with or regulates L/F transferase activity (see **1.5.**) (Soffer and Savage 1974). The elusive biological function of L/F

transferase may be further revealed by identifying interacting proteins of L/F transferase. One candidate is the chaperone GroEL protein where it has been observed to co-purify with overexpressed L/F transferase (Abramochkin and Shrader 1995).

To visualize whether L/F transferase interacts with other proteins, we have mapped conserved residues (as reported by (Suto *et al.* 2006)) on the surface of an existing crystal structure of L/F transferase in complex with a substrate analogue puromycin (PDB ID: 2DPT) (**Figure A1-1**). Visual analysis revealed surface regions displaying high conservation. The majority of the conserved surface residues were observed on the catalytic face on the central cleft between the two domains, and are heavily localized in the vicinity of the active site. Some of these residues have been shown to be crucial for the catalytic mechanism and/or substrate recognition of L/F transferase (Suto *et al.* 2006, Watanabe *et al.* 2007, Fung *et al.* 2011). Although there are few surface residues that were observed to be conserved on faces adjacent or opposite to the catalytic face, one deep surface groove leading into the catalytic site displayed heavy conservation of hydrophobic residues. This groove consists of the following residues: F95, G101, A103, I112, W135, L140, and F159. This hydrophobic groove may be responsible for L/F transferase protein-protein interactions as interactions through hydrophobic surface contacts are commonly observed. Also, L/F

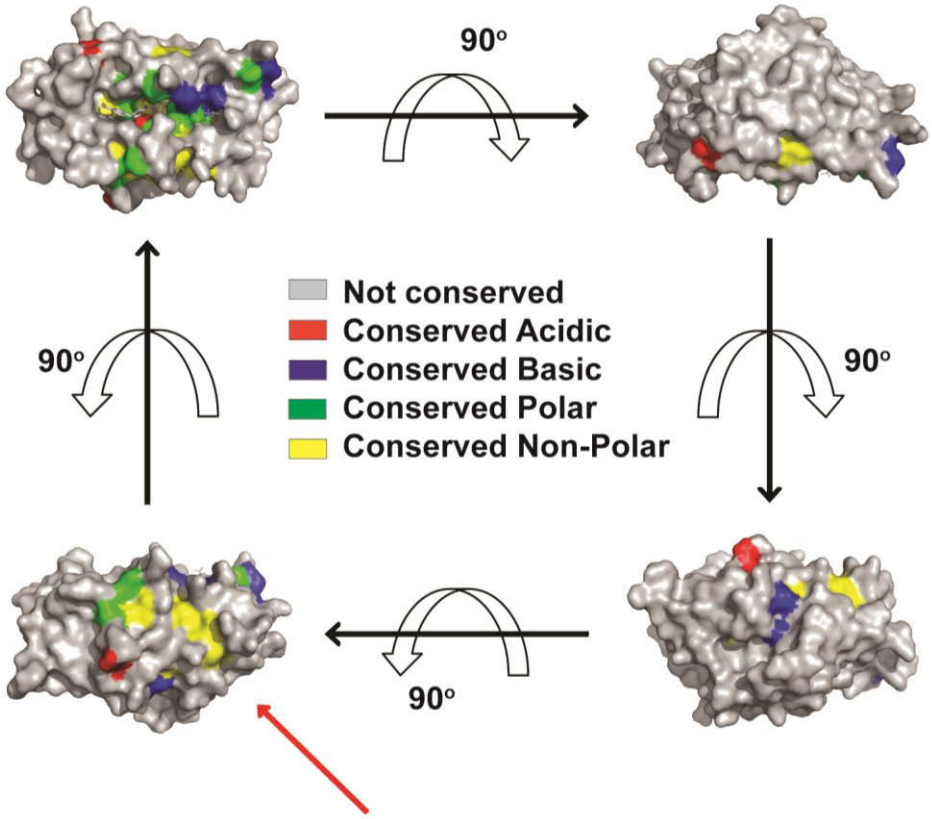

Figure A1-1: Mapping conserved residues onto L/F transferase crystal structure surface. Conserved residues organized into acidic, basic, polar, or non-polar residues from a sequence alignment **(Suto *et al.* 2006)** are mapped onto an existing L/F transferase X-ray crystal structure in complex with a substrate analogue puromycin (PDB 2DPT) using PyMoL. A deep groove with conserved hydrophobic residues exists on the surface of the protein (indicated by red arrow), suggesting a putative substrate peptide or interacting partner binding site.

transferase forming a protein-protein complex may explain the poor solubility observed (Abramochkin and Shrader 1995). In addition to protein-protein interaction, this hydrophobic groove may be responsible for substrate recognition and binding. It has been shown that N-end rule peptide substrates are required to have an unstructured linker region between the N-terminal degradation signal (N-degron) and a hydrophobic element downstream of the N-degron for efficient delivery to the proteasome-like ClpSAP protease system (Erbse *et al.* 2006, Ninnis *et al.* 2009).

During various L/F transferase protein purifications, our laboratory observed that a number of proteins consistently co-purify with L/F transferase. **Figure A1-2** shows the SDS-PAGE of GST affinity purification and size exclusion elution fractions of GST tagged L/F transferase. We have identified the chaperone DnaK and 30S ribosomal protein S1 consistently co-elutes with GST tagged L/F transferase using in-gel liquid chromatography tandem mass spectrometry (LC-MS/MS) (Institute of Biomolecular Design, University of Alberta). Other proteins identified are listed in **Table A1-1**.

Given the availability of mature protein identification methods using mass spectrometry, here we began an investigation on the putative interacting partners of L/F transferase. Utilizing affinity purification coupled to in-gel LC-MS/MS, proteins that co-purify with His-tagged and GST-tagged L/F transferase were identified. We identified the molecular

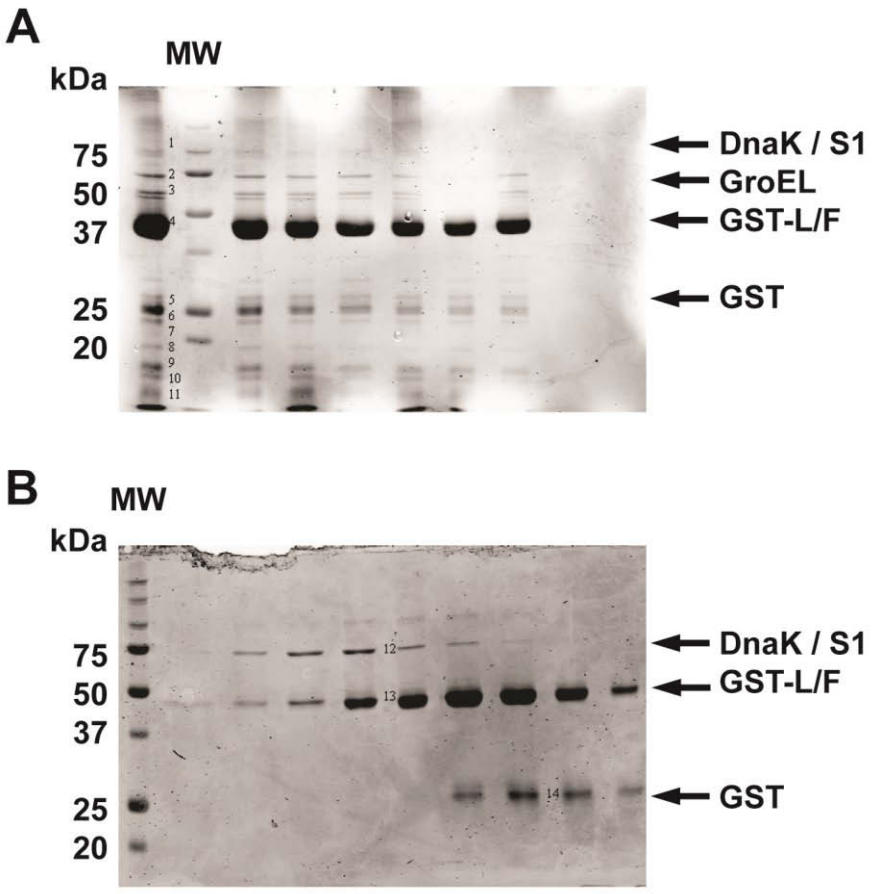

Figure A1-2: SDS-PAGE shows numerous proteins that co-purify with A) GSTrap and B) size exclusion during protein purification of GST-tagged L/F transferase. The numbered bands from elution fractions were excised and submitted for trypsin digestion and protein identification via LC-MS/MS (Institute of Biomolecular Design, University of Alberta). Arrows point to some key proteins identified, and the full identified list is listed in **Table A1-1.**

Table A1-1: List of proteins identified that co-purify with GST-L/F through GSTrap FF and HiLoad 16/60 Superdex 200 columns.

Sample #	Uniprot Accession Number	Protein	# peptide	% Coverage	MASCOT score
1	P0A6Y8 P0AG67	DnaK S1	9 7	13 12	396 344
2	P0A9D2 P0A8P1	GST L/F transferase	6 5	21 20	286
3	P0A6F5	GroEL	12	24	549
4	P0A9D2 P0A8P1	GST L/F transferase	8 6	29 25	462 266
5	P0A9D2 P0A7L0 P0A7V0	GST L1 S2	7 5 2	26 23 7	326 212 131
6	P0A9D2 P0A7V8	GST S4	8 6	29 20	402 259
7	P0A9D2 P0A7V8	GST S4	8 6	29 24	378 250
8	P62399	L5	7	34	282
9	P0ABT2 P02359	Dps S7	7 2	39 19	418 123
10	P0A7W1 P02359 P0ABT2	S5 S7 Dps	3 3 2	21 21 12	157 145 119
11	P0A7X3	S9	4	29	224
12	P0AG67 P0A6Y8	S1 DnaK	5 4	12 5	273 175
13	P0A8P1	L/F transferase	5	20	175
14	P0A9D2	GST	8	31	390

chaperonin GroEL as a potential L/F transferase interacting partner. We have additionally performed a GST pull down to validate this interaction. However, GroEL was found to be a false positive as it non-specifically interacts with glutathione agarose beads as well as GST proteins. In summary, we were unable to identify and validate any specific interacting partners for L/F transferase.

A1.2. Materials and Methods

A1.2.1. Materials

Unless stated otherwise, all chemicals were purchased from Sigma-Aldrich.

A1.2.2. Expression Vectors

A clone of the wild-type *E. coli* L/F transferase with a N-terminal hexa-histidine tag in a pCA24N expression vector was obtained from the ASKA (-) strain collection from the National Institute of Genetics (Japan). This construct will be labelled as “His-L/F”.

N-terminal Glutathione-S-Transferase-tagged L/F transferase (GST-L/F) was cloned from the above mentioned His-L/F into the pGEX-6P-1 expression vector. Primers designed with BamHI and NotI restriction sites (5'–BamHI-aat (+) primer: 5'-GGC CCC TG**G GAT CC**A TGC GCC TGG TTC AGC TT-3' and 3'-NotI-aat (-) primer: 5'- GTC ACG AT**G CGG CCG C**TC ATT CTT GTG GTG AAA ACA AGC A-3') were used in PCR to amplify the gene from the His-L/F construct. Following BamHI and NotI

restriction enzyme digest, the inserts were ligated into pGEX-6P-1 vector. pGEX-6P-1 was a generous gift from Dr. JN Mark Glover (University of Alberta). Clone was verified with DNA sequencing (The Applied Genomics Centre, Department of Medical Genetics). As a negative control, pGEX-6P-1 is also used for GST expression.

Since the *E. coli groS* (for GroES) and *groL* (for GroEL) genes were within the same operon, primers were designed with digestion sites outside the two genes for PCR amplification from whole cell genomic DNA and cloned into the pET28a (+) expression vector. Specifically the NdeI digestion site was designed 5' to the *groS* gene and the EcoRI digestion site was designed 3' to the *groL* gene (5'-NdeI-groS (+) primer: 5'- GCG GCA GC**C ATA TG**A ATA TTC GTC CAT TGC ATG-3' and 3'-EcoRI-groL (-) primer: 5'- CGG AGC TC**G AAT TC**A TTT CTG CGA GGT GCA GGG C -3'). First a wild-type K-12 colony was transferred from the LB agar plate into 20 µL of water in an eppendorf tube using a pipette tip. The tube was heated to 94 ºC for 10 min, froze to -80 ºC for 10 min, followed by centrifugation at 13 000 RPM for 5 min at 4 ºC. 5 µL of the supernatant was used as template for a typical 50 µL PCR reaction. Following NdeI and EcoRI restriction enzyme digestion, the inserts were ligated into the pET28a (+) vector. Giving an N-terminal hexa-histidine tagged GroES construct that co-purifies with GroEL, since they are within the same operon (His-GroESL). Successful clone were verified by DNA sequencing (The Applied Genomics Centre, Department of Medical Genetics).

A1.2.3. Protein Expression and Affinity Purification

In a separate experiment, GST tagged L/F transferase was affinity purified as follows. Construct GST-L/F was transformed into *E. coli* BL21 DE3 strain and grown in LB media at 37 ºC to an Abs_{600nm} of 0.4 - 0.6 prior to induction with 0.5 mM IPTG for 4 hours at 37 ºC. Cells were pelleted by centrifugation at 5 000 ×g for 10 minutes at 4 ºC. Pellets were gently resuspended in 10 mL of the Binding Buffer A (287 mM NaCl, 3 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4) supplemented with 1 mM PMSF protease inhibitor and 2 mM DTT. Cells were lysed via sonication of 2 - 4 rounds of 30 sec sonication with 5 sec pulses separated by 5 sec pauses at amplitude 70 % on ice. Cell debris was cleared by centrifugation at 10 000 RPM for 15 minutes at 4 ºC using the JA-20 rotor (Beckman Coulter). Supernatant was carefully decanted and saved.

GSTrap FF 1 mL columns (GE Healthcare) were first prepared with 5 columns of distilled water and 5 columns of Binding Buffer A using a syringe. Then lysates were applied to the column at a flow rate of 0.3 mL/min using the AKTA prime plus FPLC (GE Healthcare). Columns were subsequently washed with 10 mL of Binding Buffer A at a flow rate of 1 mL/min. Bound proteins were eluted with 10 mL of the Elution Buffer A (50 mM Tris-Cl (pH 7.4), 10 mM reduced glutathione) at a flow rate of 1 mL/min. The concentrated elution fraction (Amicon Ultra 10k MWCO Centrifugal Filter Units, Millipore) are then loaded on a pre-equilibrated size exclusion column (HiLoad 16/60 Superdex 200, GE Healhcare) with

Binding Buffer B (20 mM Tris-Cl (pH 7.4), 500mM NaCl, 1 mM DTT) for further purification.

A1.2.4. Identification of Interaction Partner via Affinity Co-Purification

To identify interaction partners of L/F transferase, constructs (His-L/F, GST-L/F and GST) were expressed as described above. Each culture was divided into two, and cells were pelleted by centrifugation at 5 000 × g for 10 minutes at 4 ºC.

HisTrap FF column Buffers: Binding Buffer C (50 mM Tris-Cl (pH 7.4), 200 mM NaCl, 10 mM KCl, 10 mM imidazole), Elution Buffer C (50 mM Tris-Cl (pH 7.4), 200 mM NaCl, 10 mM KCl, 500 mM imidazole). Note: for GroESL protein purification 10 mM $MgCl_2$ is supplemented to the buffers.

HisTrap FF and GSTrap FF 1 mL columns were washed with 5 column volumes of distilled water followed by 5 column columns of the appropriate binding buffer. Bacterial lysates from each construct (His-L/F GST, and GST-L/F) were applied to each column at flow rates of 1 drop/second using 5 mL sterile syringes. Columns were subsequently washed three times with 5 mL of the appropriate binding buffer at a flow rate of 1 drop/second using new sterile syringes. Bound proteins were eluted with 5 mL of the appropriate elution buffer at a flow rate of ½ drop/second. Elutions were collected, a final of 1 x SDS Loading Buffer were added, and stored at -20 ºC. Columns were subsequently washed

and re-generated. HisTrap FF column was first stripped with 5 volumes of 50 mM EDTA, followed by distilled water, and stored in 20% EtOH. GSTrap FF column was first stripped with 2 volumes of 6 M guanidine hydrochloride, followed by 5 volumes of Binding Buffer A, 70% EtOH, Binding Buffer A, and stored in 20% EtOH.

A1.2.5. SDS-PAGE and LC-MS/MS Protein Identification

Affinity column elutions were visualized via 10% SDS-PAGE to reduce sample complexity prior to liquid chromatography tandem mass spectrometry. Gels were stained overnight in Coomassie Brilliant Blue G-250 protein stain. SDS-PAGE gels were destained, and each lane was cut into approximately 10 gel slices. Each gel slice was further cut into 1 mm x 1 mm gel fragments, and placed into a 96-well micro-titer plates.

The micro-titer plate was placed into a PerkinElmer Mass PREP-Station digest-robot, where gel fragments were reduced with 10 mM β-mercaptoethanol, alkylated with 55 mM iodoacetamide, dehydrated in acetonitrile, and digested with 6 ng/μL trypsin in 100 mM ammonium bicarbonate. Tryptic peptides were then extracted in 1% formic acid and 2% acetonitrile and separated by reverse-phase HPLC by running a 40 minute reverse-phase acetonitrile/water gradient over two 280 mm / 50 μm Agilent C18 trapping columns. HPLC elutions were fed into a Thermo-Finnigan LCQ DecaXP ion trap mass spectrometer or Thermo LTQ XL Orbitrap mass spectrometer via electrospray ionization, and subject to tandem mass spectrometry analysis. Mass spectra were collected using

the program XCalibur. Mass spectra were subsequently searched against the SwissProt *E. coli* protein database using the MASCOT search engine.

A1.2.6. Immunoprecipitation of GST-L/F with His-GroESL

Proteins His-GroESL, GST-L/F, and GST were each affinity purified as described above except with using the AKTA prime plus FPLC (GE Healthcare) instead of syringes. GST-L/F and GST proteins were purified using the GSTrap FF column, and elutions were dialyzed into the Dialysis Buffer A (20 mM Tris-Cl (pH 7.4) 500 mM NaCl, 1 mM DTT, 10 % glycerol). His-GroESL was purified using the HisTRap FF column, and elutions were dialyzed into Dialysis Buffer B (50 mM Tris-Cl (pH 7.4), 200 mM NaCl, 2 mM $MgCl_2$, 1 mM DTT, 10 % glycerol). However, His-GroESL protein precipitated during dialysis. For subsequent purifications, elution fractions were pooled and centrifuged and split to dialyzed in a high salt buffer (20 mM Tris-Cl (pH 7.4), 1 M NaCl, 10 % glycerol) or low salt buffer (20 mM Tris-Cl (pH 7.4), 200 mM NaCl, 10 % glycerol). His-GroESL remain soluble in both high salt buffer and low salt buffer.

For the GST immunoprecipitation, 60 µL of 60% mixed slurry glutathione-agarose beads were added to three eppendorf tubes. The beads were centrifuged at 6 000 RPM for 10 min at 4 ºC to remove excess buffer. The beads were washed with Binding Buffer D (137 mM NaCl, 3 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4) for three times. The beads were resuspended to a 50 % slurry by adding Binding Buffer D. Incubate the beads with bait protein (buffer control, 0.15 µg of GST control, or 0.15

μg of GST-L/F) at 4 ºC with agitation for 2 hours. The beads were centrifuged to remove supernatant, and an aliquot was saved as bait input. The beads were washed with 1 mL Binding Buffer D for three times and incubated with prey protein (0.6 μg of His-GroESL) overnight at 4 ºC with agitation. The beads were centrifuged to remove supernatant, and an aliquot was saved as prey input. The beads were washed with 1 mL Binding Buffer D for three times and incubated with 60 μL of Elution Buffer D (20 mM Tris-Cl (pH 7.4), 50 mM NaCl, 10 mM reduced glutathione) for 30 min at room temperature with agitation. The beads were centrifuged to remove supernatant, and an aliquot was saved as elution. Each sample were then added to a final of 1 x SDS Loading Buffer, and analysed by 14 % SDS-PAGE.

A1.3. Results and Discussion

A1.3.1. The Identification of Potential Interacting Partners via Co-Purification

To identify potential interacting partners of L/F transferase, we employed the affinity purification coupled with in-gel LC-MS/MS (AP-MS) method. Since Ni^{2+} affinity columns are known for non-specifically pulling down proteins that lack His-affinity tags (Lichty *et al.* 2005), a number of negative controls are considered and employed to eliminate the reporting of false positives. **Figure A1-3** summarizes the affinity purification design. We have two affinity tagged L/F transferase constructs, hexa-histidine tagged (His-L/F) and GST-tagged (GST-L/F). Each construct's lysate will

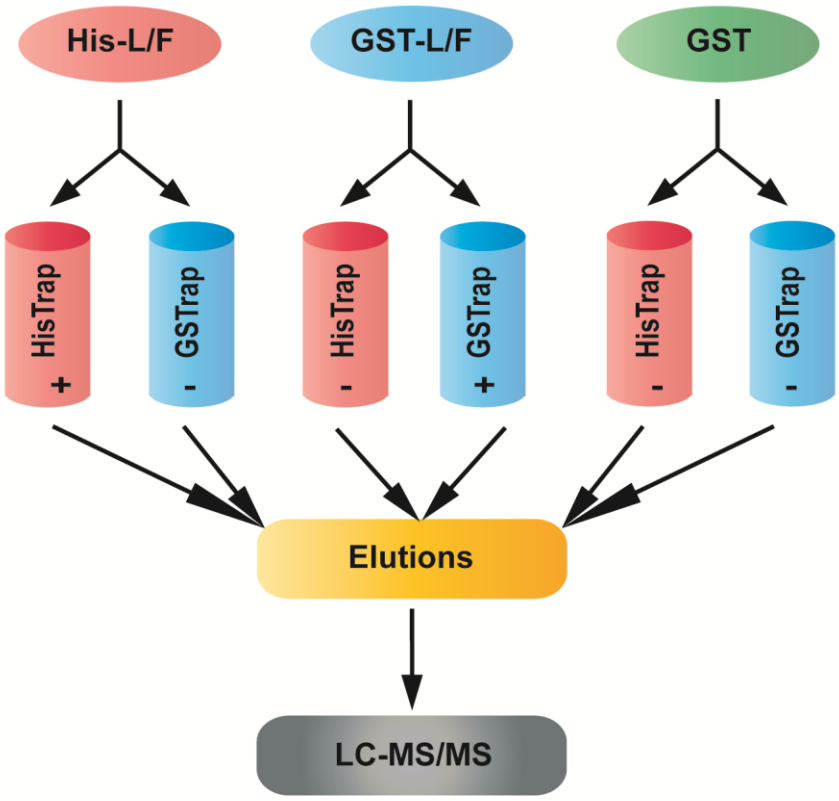

Figure A1-3: Experimental design for the identification of co-purifying proteins. BL21 DE3 cells expressing His-L/F, GST-L/F and GST were divided into two and lysed. Each construct was resuspended in HisTrap Binding Buffer B or GSTrap Binding Buffer A, and subject to either HisTrap or GSTrap affinity purification. Elutions were collected for SDS-PAGE analysis and LC-MS/MS protein identification. (+) sign indicates experiment and (-) sign indicates negative control.

be divided evenly and applied to both HisTrap FF and GSTrap FF columns (GE Healthcare). The negative controls include the His-L/F elution fraction from the GSTrap column, GST-L/F elution fraction from the HisTrap column, and GST elution fractions from HisTrap and GSTrap columns.

After SDS-PAGE analysis of the affinity purification elution fractions, visual analysis of the gels indicated a large number of false positives being pulled down in the HisTrap columns, while the GSTrap column maintained specificity (**Figure A1-4**). Subsequent LC-MS/MS (ion trap) (**Figure A1-5**) identified 30 proteins from His-LF (HisTrap) elutions, 40 proteins from GST-LF (HisTrap) elutions, 12 proteins from GST (HisTrap) elutions, 5 proteins eluting from GST-LF (GSTrap) elutions, 0 proteins from His-LF (GSTrap) elutions, and 3 proteins from GST (GSTrap) elutions. Later we also re-analyzed the samples by the new Thermo Orbitrap LTQ XL mass spectrometer, where we identified 47 proteins from His-LF (HisTrap) elutions, 66 proteins from GST-LF (HisTrap) elutions, 6 proteins from GST (HisTrap) elutions, 8 proteins eluting from GST-LF (GSTrap) elutions, 1 protein from His-LF (GSTrap) elutions, and 2 proteins from GST (GSTrap) elutions. **Figure A1-6** summarizes the findings in a Venn diagram. L/F transferase was positively identified in His-LF (HisTrap) elutions, indicating successful expression and pull-down of our protein of interest. In addition, glutathione-S-transferase and L/F transferase were identified in GST-LF

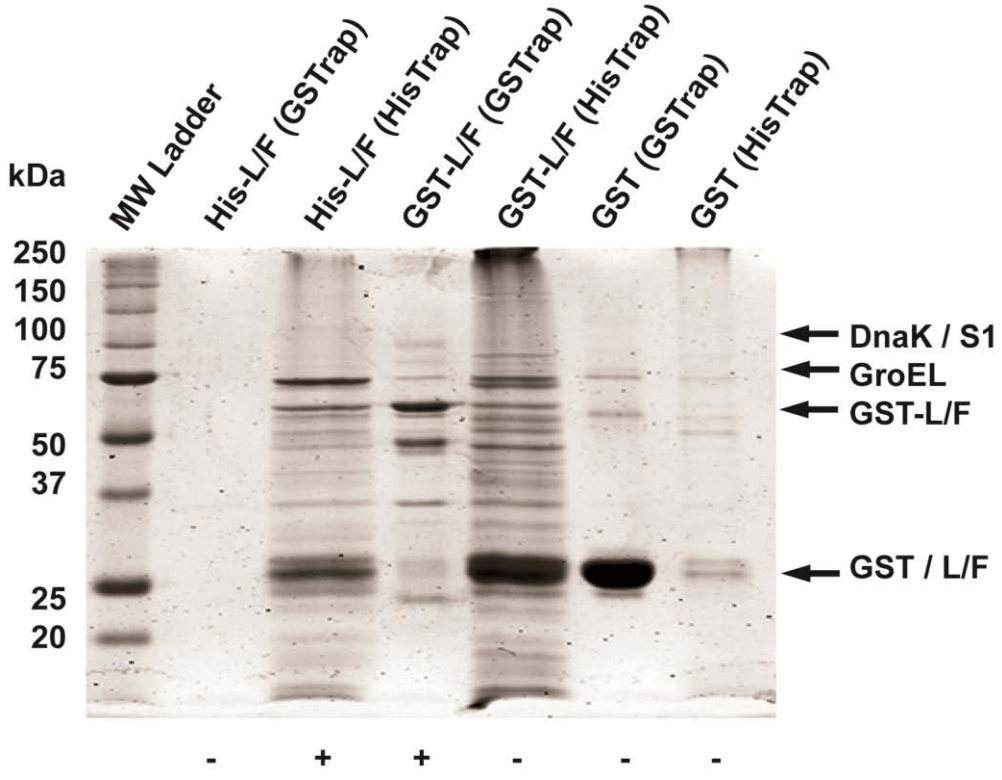

Figure A1-4: SDS-PAGE analysis of HisTrap and GSTrap elution fractions for His-L/F and GST-L/F, and GST. HisTrap columns overall contains more non-specific proteins than GSTrap columns. Arrows point to some key proteins identified and the full lists are listed in **Table A1-2** and **Table A1-3.**

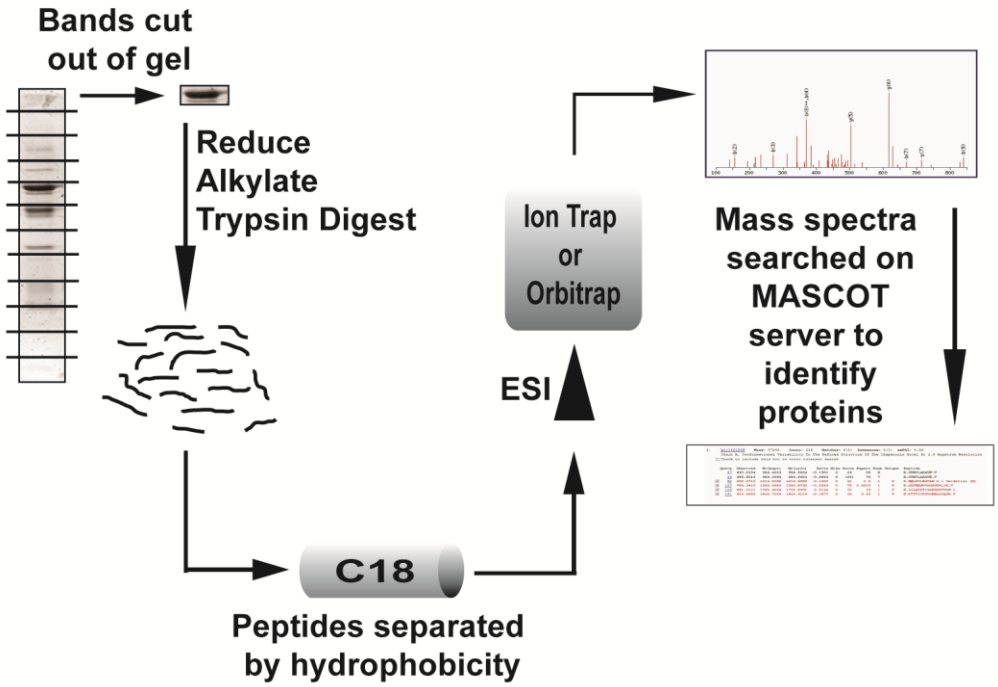

Figure A1-5: Schematic for LC-MS/MS protein identification. Each elution lane was excised and separate into diced bands. Each gel slice undergoes reduction and alkylation to deactivate cysteine sulfhydryl groups, and trypsin digestion. Digested peptides were separated by hydrophobicity via C18 column HPLC, and feed into an Ion Trap or Orbitrap mass spectrometer via electrospray ionization. Mass spectra were collected and searched by MASCOT for protein identification.

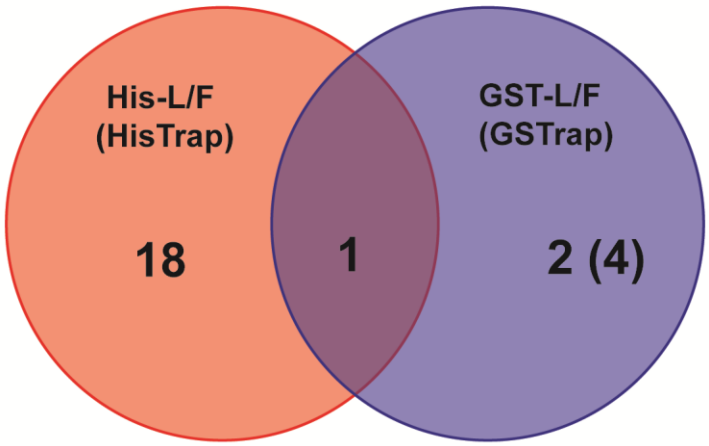

Figure A1-6: Venn diagram comparison of identified co-purifying proteins. The diagram lists the number of proteins identified by ion trap after subtracting the negative controls. The number in the bracket indicates the number of proteins identified by orbitrap as a replicate. A single protein – 60kDa chaperone GroEL - is identified to be common and co-purifies with the two affinity tagged L/F transferase.

(HisTrap) and GST-LF (GSTrap) column elutions, indicating successful fusion of GST with L/F transferase.

Tables A1-2 and A1-3 summarize the lists of proteins identified to both experiment columns (His-LF (HisTrap) and GST-LF (GSTrap)) where the negative control proteins identified have been removed (GST-LF and GST (HisTrap) and His-L/F and GST (GSTrap) respectively). A single protein, the molecular chaperone GroEL, was common between both experiment columns. Although molecular chaperones are often common contaminants in protein purifications, GroEL was thought not to be background because a previous pull down performed for a different bacterial protein RelA (data not shown) following the same procedure did not lead to the identification of GroEL. Additionally, this observation is consistent with a previous identification of the chaperonin GroEL to bind to Ni-NTA agarose column only in the presence of L/F transferase (Abramochkin and Shrader 1995).

The ~60 kDa GroEL, a member of the chaperonin family of molecular chaperones, is responsible for assisting in the proper folding of proteins which have either remained unfolded due to non-permissive folding conditions or have misfolded upon ribosomal release (Goldberg 2003, Ellis 2005, Kerner *et al.* 2005, Azia *et al.* 2012). Misfolded proteins may aggregate due to the presence of hydrophobic residues exposed on their surface. GroEL-dependent protein folding plays a crucial role in cell-survival. The functional role for molecular chaperones in protein

Table A1-2: List of proteins identified that co-purify with His-L/F through HisTrap FF column (negative controls subtracted).

Uniprot Accession Number	Protein	# peptide	% Coverage	MASCOT score
Ion Trap				
P0A6F5	GroEL	11	30	518
P02359	30S ribosomal protein S7	6	48	309
P0A8P1	L/F transferase	7	32	268
P0AA10	50S ribosomal protein L13	5	48	193
B1XBS1	bifunctional chorismate mutase/prephenate dehydratase	4	9	183
P02413	50S ribosomal protein L15	4	35	179
P0A7R1	50S ribosomal protein L9	2	18	133
P0A7B8	Chain A, Hslv (Clpq)	2	16	132
P0A9A9	ferric uptake regulator	3	27	127
P61175	50S ribosomal protein L22	2	19	121
P0A7X3	30S ribosomal protein S9	3	21	120
P0AG55	50S ribosomal protein L6	3	22	107
P39208	gluconate kinase 1	2	11	95

P60438	50S ribosomal protein L3	3	21	94
P62399	50S ribosomal protein L5	2	17	90
P0ADY7	50S ribosomal protein L16	2	22	89
P21513	ribonuclease III	3	11	85
P0A7W1	30S ribosomal protein S5	2	15	84
	O137	2	15	81
P0A7V0	30S ribosomal protein S2	2	11	58
Orbitrap				
P0A8P1	L/F transferase	8	54	773
P0A6F5	GroEL	9	36	165
P17169	D-fructose-6-phosphate amidotransferase	6	18	115
P0A7L0	50S ribosomal protein L1	6	21	69
P0A717	Ribose-phosphate pyrophosphokinase	4	19	39
	Component in Transcription anti-termination	4	28	39
P60716	Lipoate synthesis protein LipA	2	12	36
P00484	Chloramphenicol acetyl transferase	3	16	33
P62399	50S ribosomal protein L5	2	18	29

P0A707	Translation IF-3	2	18	24
P69795	PTS enzyme II AB	2	10	22
P0ADG4	Inositol monophosphatase	2	11	22
P0ABA0	Membrane bound ATP synthase, F1 sector, beta subunit	3	9	18
P0A6Q3	fabA	2	16	16
P04805	Glutamyl-tRNA synthetase	2	5	15
P37773	UDP-N-acetylmuramate: L-alanyl-gamma-D-glutamyl-meso-diaminopimelate ligase	2	9	14
P39280	Lysine aminomutase	2	10	14
P0A8P8	Site-specific Tyrosine recombinase XerD	2	8	12
P23908	Acetylornithine deacetylase	2	10	11
	Putative solute-DNA competence effector	2	13	10

Table A1-3: List of proteins identified to co-purify with GST-L/F through GSTrap FF column (negative controls subtracted).

Uniprot Accession Number	Protein	# peptide	% Coverage	MASCOT score
Ion Trap				
P0A6F5	GroEL	9	21	393
P0A8P1	L/F transferase	3	16	120
P0CE47	Chain A, Elongation Factor Complex EF-Tu:EF-Ts	2	5	95
Orbitrap				
P0A8P1	L/F transferase	8	50	394
P0A6F5	GroEL	12	45	335
P06959	Dihydrolipoamide acetyltransferase	4	11	48
P0A6Y8	DnaK	6	11	38
P60723	50S ribosomal protein L4	2	17	21

degradation has been demonstrated previously (Straus *et al.* 1988, Sherman and Goldberg 1992), and perhaps suggest a functional role of GroEL in the N-end rule Pathway. Or perhaps GroEL may be required for maintaining the solubility of affinity tagged L/F transferase.

A1.3.2. GroESL is not an Interaction Partner of L/F transferase via in vitro GST Immunoprecipitation

Since the chaperonin GroEL consistently co-purifies with both His-tagged and GST-tagged L/F transferase, we performed an *in vitro* GST immunoprecipitation experiment to validate this interaction. First, I designed primers with restriction sites just outside the *groS* and *groL* genes for whole cell genome PCR amplification. After restriction digestion, the amplified fragment is cloned into the expression vector pET28a (+). Upon IPTG expression, a N-terminal hexa-histidine tagged GroES is expressed. GroEL is also expressed under the control of the same promoter, and can be purified together as the GroESL complex.

His-GroESL was purified using the standard Ni^{2+} affinity purification, and dialyzed into the appropriate buffers. GST-L/F and GST were also purified using the GSTrap FF column, and dialyzed into the appropriate buffers. However, His-GroESL precipitated out of solution during dialysis (data not shown), despite testing twelve different buffers for solubility. Thus the preliminary GST pull down was performed immediately following HisTrap affinity purification of His-GroESL. **Figure A1-7** shows the SDS-PAGE analysis of GST immunoprecipitation using glutathione agarose

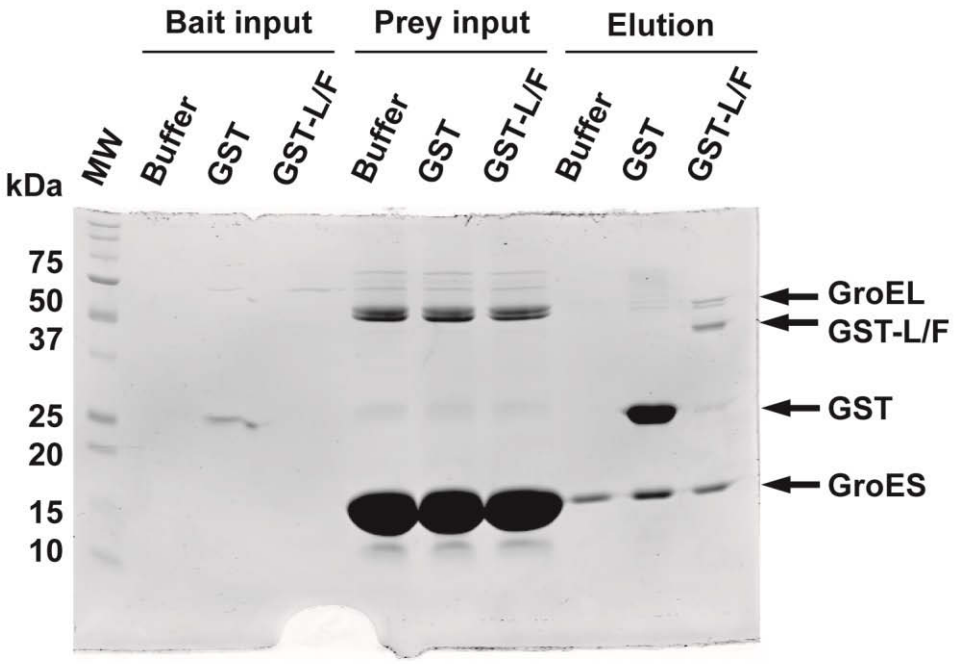

Figure A1-7: SDS-PAGE of GST pull down of GroESL (from pooled elution fractions). His-GroESL co-elutes with glutathione beads alone (buffer), GST, and GST-L/F.

beads with binding buffer, GST, or GST-L/F as bait and His-GroESL (from HisTrap FF elution fraction) as prey. The prey input lanes clearly shows the presence of His-GroESL. The elution shows that His-GroESL binds to glutathione agarose beads non-specifically as it co-elutes with buffer control, GST control, as well as GST-L/F.

Meanwhile, another His-GroESL affinity purification was performed. The HisTrap elution fractions were centrifuged and the soluble supernatant were dialyzed into either a high salt or low salt buffer. A GST pull down was performed similarly as described above, and the SDS-PAGE analysis is shown in **Figure A1-8**. Similar to the above observation, we found that His-GroESL binds to glutathione agarose beads non-specifically under both high and low salt buffer conditions. We concluded that the identification of GroEL as an interacting partner of L/F transferase was a false positive.

A1.3.3. Concluding Remarks

Through the use of affinity purification of His-tagged and GST-tagged L/F transferase coupled to in-gel LC-MS/MS, a single protein – molecular chaperone GroEL – was identified as the first potential L/F transferase interacting partner. However, GroEL was then proven to be a false positive using *in vitro* GST immunoprecipitation methods. Thus, we were unable to identify any specific interacting partners for L/F transferase.

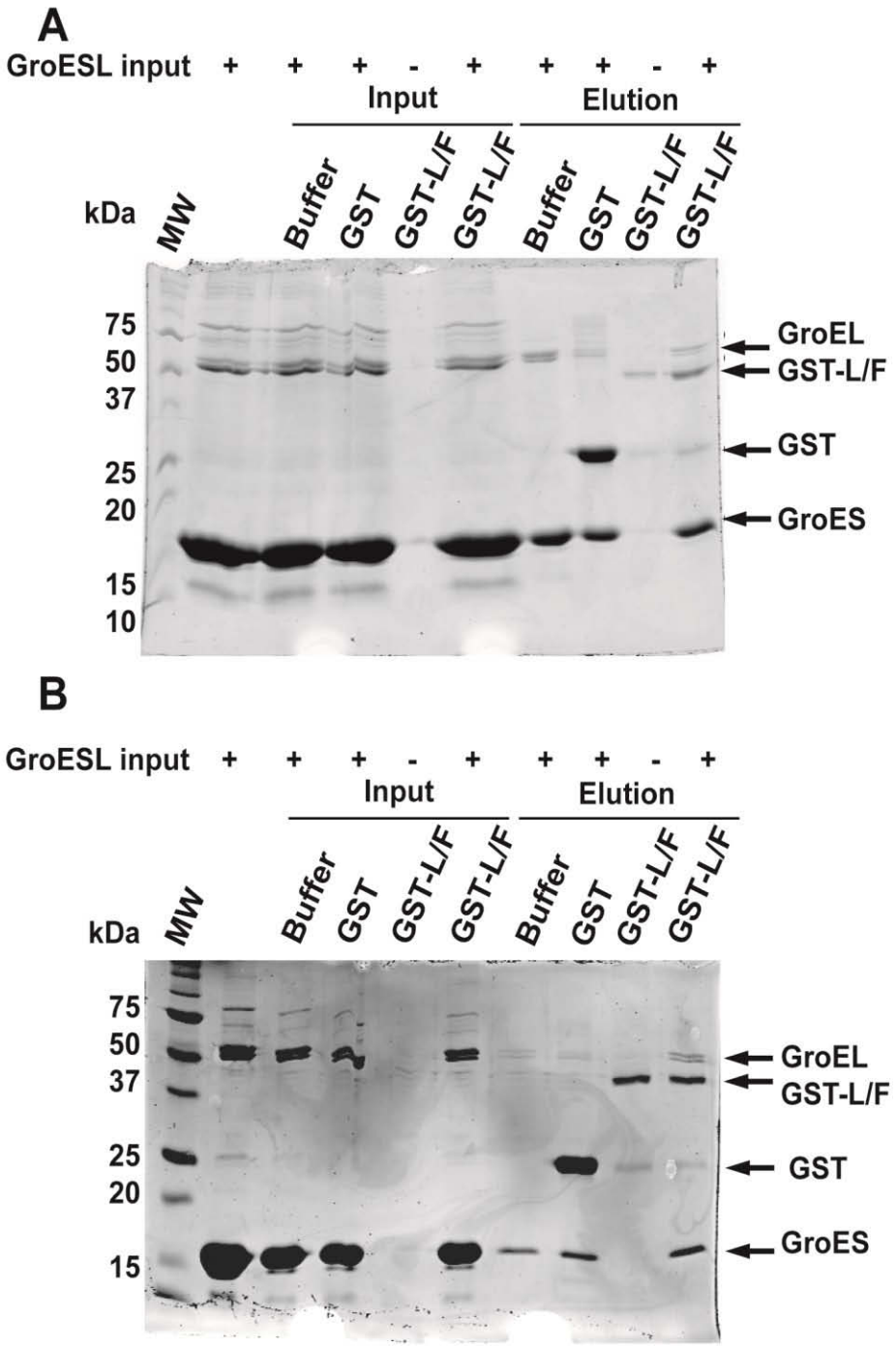

Figure A1-8: SDS-PAGE of GST pull down of GroESL (in A) high or B) low salt buffer). His-GroESL co-elutes with glutathione beads alone (buffer), GST, and GST-L/F under both high and low salt buffer. This suggests that His-GroESL does not interact with L/F transferase.

More recently two high throughput investigations that map interacting proteins of *E. coli* have used L/F transferase as a bait protein to identify protein-protein interactions (Arifuzzaman *et al.* 2006, Rajagopala *et al.* 2014). Arifuzzaman *et al.* performed a large scale pull down using the His-tagged ORF library (4,339 bait proteins tested (Kitagawa *et al.* 2005)) and Ni^{2+}-NTA beads to identify interacting proteins by MALDI-ToF MS (Arifuzzaman *et al.* 2006). Rajagopala *et al*. performed a large scale binary yeast two-hybrid screen (3,305 bait proteins, ~70 % of the *E. coli* proteome) for protein-protein interactions (Rajagopala *et al.* 2014). The putative interacting partners of L/F transferase are listed in **Table 1-1**, but have not been validated. Many putative interacting partners of L/F transferase belong to protein complexes and have wide biological functions including DNA replication, translation, and metabolism. Interestingly Arifuzzaman *et al.* also identified GroEL as a putative interacting partner despite the fact that it is also present in the control experiment (Arifuzzaman *et al.* 2006). This suggests and confirms that GroEL is not likely an interacting partner of L/F transferase.

The functional role of the deep hydrophobic groove leading to the active site of L/F transferase remains elusive. Future experiments that identify additional *in vivo* protein substrates of L/F transferase may aid in characterizing the relationship between the hydrophobic groove and protein substrate specificity or protein-protein interactions.

A1.4. References

Abramochkin, G., and Shrader, T.E. (1995) The leucyl/phenylalanyl-tRNA-protein transferase. Overexpression and characterization of substrate recognition, domain structure, and secondary structure*. J.Biol.Chem.* **270,** 20621-20628

Arifuzzaman, M., Maeda, M., Itoh, A., Nishikata, K., Takita, C., Saito, R., Ara, T., Nakahigashi, K., Huang, H.C., Hirai, A., Tsuzuki, K., Nakamura, S., Altaf-Ul-Amin, M., Oshima, T., Baba, T., Yamamoto, N., Kawamura, T., Ioka-Nakamichi, T., Kitagawa, M., Tomita, M., Kanaya, S., Wada, C., and Mori, H. (2006) Large-scale identification of protein-protein interaction of *Escherichia coli* K-12*. Genome Res.* **16,** 686-691

Azia, A., Unger, R., and Horovitz, A. (2012) What distinguishes GroEL substrates from other *Escherichia coli* proteins?*. FEBS J.* **279,** 543-550

Dong, X., Kato-Murayama, M., Muramatsu, T., Mori, H., Shirouzu, M., Bessho, Y., and Yokoyama, S. (2007) The crystal structure of leucyl/phenylalanyl-tRNA-protein transferase from *Escherichia coli. Protein Sci.* **16,** 528-534

Ebhardt, H.A., Xu, Z., Fung, A.W., and Fahlman, R.P. (2009) Quantification of the post-translational addition of amino acids to proteins by MALDI-TOF mass spectrometry*. Anal.Chem.* **81,** 1937-1943

Ellis, R.J. (2005) Chaperomics: *in vivo* GroEL function defined*. Curr.Biol.* **15,** R661-3

Erbse, A., Schmidt, R., Bornemann, T., Schneider-Mergener, J., Mogk, A., Zahn, R., Dougan, D.A., and Bukau, B. (2006) ClpS is an essential component of the N-end rule pathway in *Escherichia coli. Nature.* **439,** 753-756

Fung, A.W., Ebhardt, H.A., Abeysundara, H., Moore, J., Xu, Z., and Fahlman, R.P. (2011) An alternative mechanism for the catalysis of peptide bond formation by L/F transferase: substrate binding and orientation*. J.Mol.Biol.* **409,** 617-629

Fung, A.W., Ebhardt, H.A., Krishnakumar, K.S., Moore, J., Xu, Z., Strazewski, P., and Fahlman, R.P. (2014a) Probing the Leucyl/Phenylalanyl tRNA Protein Transferase Active Site with tRNA Substrate Analogues*. Protein Pept.Lett.* **21,** 603-614

Fung, A.W., Leung, C.C., and Fahlman, R.P. (2014b) The determination of $tRNA^{Leu}$ recognition nucleotides for *Escherichia coli* L/F transferase*. RNA.* **20,** 1210-1222

Goldberg, A.L. (2003) Protein degradation and protection against misfolded or damaged proteins*. Nature.* **426,** 895-899

Humbard, M.A., Surkov, S., De Donatis, G.M., Jenkins, L.M., and Maurizi, M.R. (2013) The N-degradome of *Escherichia coli*: limited proteolysis *in vivo* generates a large pool of proteins bearing N-degrons*. J.Biol.Chem.* **288,** 28913-28924

Kawaguchi, J., Maejima, K., Kuroiwa, H., and Taki, M. (2013) Kinetic analysis of the leucyl/phenylalanyl-tRNA-protein transferase with acceptor peptides possessing different N-terminal penultimate residues*. FEBS Open Bio.* **3,** 252-255

Kerner, M.J., Naylor, D.J., Ishihama, Y., Maier, T., Chang, H.C., Stines, A.P., Georgopoulos, C., Frishman, D., Hayer-Hartl, M., Mann, M., and Hartl, F.U. (2005) Proteome-wide analysis of chaperonin-dependent protein folding in *Escherichia coli. Cell.* **122,** 209-220

Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., and Mori, H. (2005) Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research*. DNA Res.* **12,** 291-299

Leibowitz, M.J., and Soffer, R.L. (1969) A soluble enzyme from *Escherichia coli* which catalyzes the transfer of leucine and phenylalanine from tRNA to acceptor proteins*. Biochem.Biophys.Res.Commun.* **36,** 47-53

Lichty, J.J., Malecki, J.L., Agnew, H.D., Michelson-Horowitz, D.J., and Tan, S. (2005) Comparison of affinity tags for protein purification*. Protein Expr.Purif.* **41,** 98-105

Ninnis, R.L., Spall, S.K., Talbo, G.H., Truscott, K.N., and Dougan, D.A. (2009) Modification of PATase by L/F-transferase generates a ClpS-dependent N-end rule substrate in *Escherichia coli. EMBO J.* **28,** 1732-1744

Rajagopala, S.V., Sikorski, P., Kumar, A., Mosca, R., Vlasblom, J., Arnold, R., Franca-Koh, J., Pakala, S.B., Phanse, S., Ceol, A., Hauser, R., Siszler, G., Wuchty, S., Emili, A., Babu, M., Aloy, P., Pieper, R., and Uetz, P. (2014) The binary protein-protein interaction landscape of *Escherichia coli. Nat.Biotechnol.* **32,** 285-290

Scarpulla, R.C., Deutch, C.E., and Soffer, R.L. (1976) Transfer of methionyl residues by leucyl, phenylalanyl-tRNA-protein transferase. *Biochem.Biophys.Res.Commun.* **71,** 584-589

Schmidt, R., Zahn, R., Bukau, B., and Mogk, A. (2009) ClpS is the recognition component for *Escherichia coli* substrates of the N-end rule degradation pathway. *Mol.Microbiol.* **72,** 506-517

Sherman, M.Y., and Goldberg, A.L. (1992) Heat shock in *Escherichia coli* alters the protein-binding properties of the chaperonin groEL by inducing its phosphorylation. *Nature.* **357,** 167-169

Shrader, T.E., Tobias, J.W., and Varshavsky, A. (1993) The N-end rule in *Escherichia coli*: cloning and analysis of the leucyl, phenylalanyl-tRNA-protein transferase gene *aat. J.Bacteriol.* **175,** 4364-4374

Soffer, R.L. (1973) Peptide acceptors in the leucine, phenylalanine transfer reaction. *J.Biol.Chem.* **248,** 8424-8428

Soffer, R.L., and Savage, M. (1974) A mutant of *Escherichia coli* defective in leucyl, phenylalanyl-tRNA-protein transferase. *Proc.Natl.Acad.Sci.U.S.A.* **71,** 1004-1007

Straus, D.B., Walter, W.A., and Gross, C.A. (1988) *Escherichia coli* heat shock gene mutants are defective in proteolysis. *Genes Dev.* **2,** 1851-1858

Suto, K., Shimizu, Y., Watanabe, K., Ueda, T., Fukai, S., Nureki, O., and Tomita, K. (2006) Crystal structures of leucyl/phenylalanyl-tRNA-protein transferase and its complex with an aminoacyl-tRNA analog. *EMBO J.* **25,** 5942-5950

Wagner, A.M., Fegley, M.W., Warner, J.B., Grindley, C.L., Marotta, N.P., and Petersson, E.J. (2011) N-terminal protein modification using simple aminoacyl transferase substrates. *J.Am.Chem.Soc.* **133,** 15139-15147

Watanabe, K., Toh, Y., Suto, K., Shimizu, Y., Oka, N., Wada, T., and Tomita, K. (2007) Protein-based peptide-bond formation by aminoacyl-tRNA protein transferase. *Nature.* **449,** 867-871

Appendix 2

The Identification of L/F transferase *in vivo* Substrates by Click Chemistry Coupled in-gel LC-MS/MS

Note: Zhizhong Xu (Susan) contributed to the mutagenesis of PheRS. Michal Gozdzik (Summer Student in 2012 and 2013) contributed to the immunoblot, fractionation and optimization of streptavidin pull down procedures.

A2.1. Introduction

The biological functions of L/F transferase and N-end rule in *E. coli* remain enigmatic is partly due to the lack of characterization of its *in vivo* substrates. Preliminary attempts in identifying putative L/F transferase substrates found at least 21 soluble and 3 ribosomal acceptor proteins via the acylation of Δ*aat* lysates with [^{14}C]-phenylalanine after the addition of purified L/F transferase (Leibowitz and Soffer 1971b, Soffer and Savage 1974). Since that study, most L/F transferase studies focused on model substrates and mechanisms.

Recently with the identification that the adaptor protein ClpS modulates the substrate specificity for the ClpAP protease complex (Erbse *et al.* 2006), it has been rationalized that proteins that interact with ClpS are N-end rule substrates. Two studies have identified ClpS-interacting proteins (Ninnis *et al.* 2009, Schmidt *et al.* 2009). Of the twenty or so identified ClpS-interacting substrates, only two are confirmed as N-end rule substrates: DNA protection during starvation (Dps) and putrescine aminotransferase (PATase) (see **1.5.**) (Ninnis *et al.* 2009, Schmidt *et al.* 2009).

L/F transferase has become a recent focus for the development of protein engineering. The first use of unnatural amino acids by L/F transferase was demonstrate over 40 years ago using *p*-fluorophenylalanine (Leibowitz and Soffer 1969, Rao and Kaji 1974). Since then L/F transferase has been used to conjugate chemically diverse

unnatural amino acids to the N-termini of peptides and proteins and further coupled the labelled peptides and proteins to specific biotin or fluorescent probes, resulting in the formation of artificially tagged peptides and proteins (Kuno *et al.* 2003, Taki and Sisido 2007, Connor *et al.* 2008, Taki *et al.* 2008, Ebisu *et al.* 2009, Taki *et al.* 2009, Wagner *et al.* 2011). This novel technique may be applied to a variety of applications, such as enriching peptides from complex cell lysate mixtures in proteomics studies.

The specific goal of this study is to identify the *in vivo* substrates of L/F transferase using a targeted proteomic approach. My strategy is to use the enzymatic activity of L/F transferase to selectively label its substrates with unnatural amino acids to facilitate their isolation and subsequent identification. Of all the tested unnatural amino acids, we are particularly interested in *p*-azido-phenylalanine (azido-Phe) because of its capability to react with alkyne groups via copper (I)-catalyzed [3+2] cycloaddition (click chemistry). Click chemistry is a highly selective, specific and rapid chemical reaction that is compatible with biomolecules (Rostovtsev *et al.* 2002, Tornoe *et al.* 2002, Wang *et al.* 2003, Speers and Cravatt 2004). The identification of *in vivo* substrates is essential for the comprehension of the biological roles of L/F transferase.

Figure A2-1 shows the experimental outline for the identification of *in vivo* L/F transferase by click chemistry coupled LC-MS/MS. The use of Δ*aat* cell lysates ensures that potential substrates have not already been

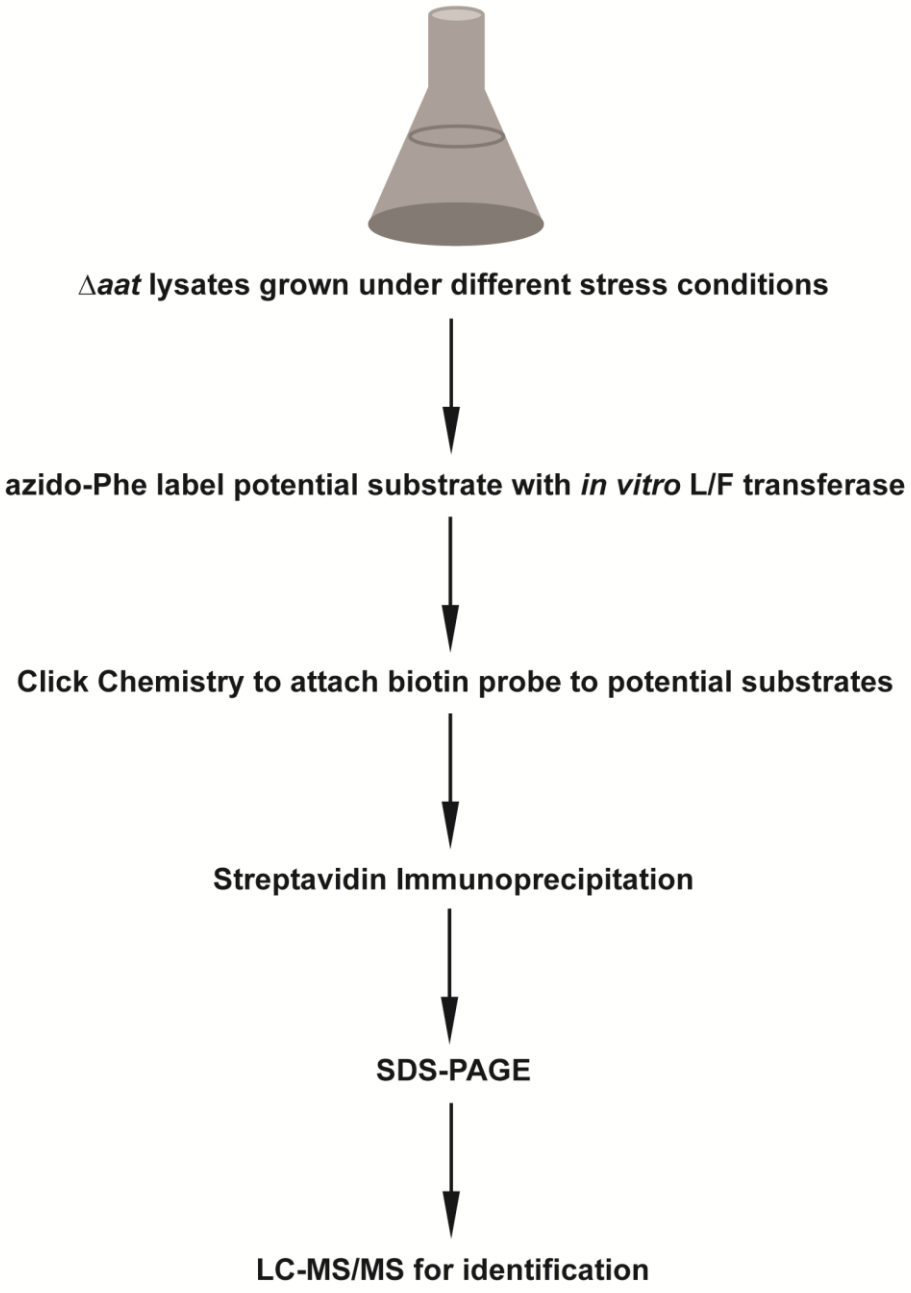

Figure A2-1: Experimental Outline for the identification of L/F transferase *in vivo* substrates by click chemistry coupled LC-MS/MS.

modified by endogenous L/F transferase. We aim to label potential protein substrates with the unnatural amino acid azido-Phe. For detection, the azido-Phe labelled proteins can be modified with a fluorescent or biotin alkyne probe using the click chemistry procedure. A fluorescent alkyne probe such as NBD-alkyne enables rapid detection and quantification of the modified proteins after being resolved by SDS-PAGE. A biotin-alkyne probe enables enrichment procedures via immunoprecipitation with avidin agarose beads. Biotin-Avidin interaction is an extremely strong non-covalent interaction with a K_a of 10^{15} M^{-1} (Green 1975). The labeled and derivatized samples will be resolved by gel electrophoresis and the entire gel lane will be excised and treated to an in-gel trypsin digestion procedure (**A1.2.5**). The eluted peptides will then be analyzed by LC-MS/MS on the in-house LTQ Orbitrap mass spectrometer (Thermo Scientific). Proteins will be identified by analyzing the data by SEQUEST (Thermo Scientific).

Here we demonstrate the optimization steps prior to the identification of *in vivo* L/F transferase protein substrates. First, the azido-Phe labelling and click chemistry procedures were optimized using model peptide substrates to ensure the feasibility of the chemical reactions. Secondly, the isolation procedure depends heavily on biotinylation. There is a single known biotinylated protein in *E. coli*, the biotin carboxyl carrier protein (BCCP) subunit of acetyl-CoA carboxylase (~17kDa) (Chapman-Smith and Cronan 1999), but the presence of other could result in false

positive identifications. We demonstrated and identified that there are a variety of low abundant biotinylated proteins in *E. coli*. Finally, a significant challenge to this aim is the lack of knowledge in the cellular events responsible for the generation of L/F transferase substrates. To overcome this challenge, we prepared the lysates from *⊿aat* cells cultured under several temperature and nutrient stress conditions that are known to induce a proteolytic cascade. However, we were unable to identify the stress event that induce N-end rule in *E. coli*. This documents some key initial optimization and experiments towards the identification of *in vivo* L/F transferase protein substrates.

A2.2. Materials and Methods

A2.2.1. Materials

Unless stated otherwise, all chemicals were purchased from Sigma-Aldrich. Peptides REPGLCTWQSLR and FREPGLCTWQSLR were purchased from Institute of Biomolecular Design (University of Alberta). The unnatural amino acid *p*-azidophenylalanine (azido-Phe) was purchased from Chem-Impex International Inc.

A2.2.2. Expression Vectors and Protein Purification

A clone of a 6× histidine tagged *E. coli* phenylalanyl-tRNA synthetase (PheRS) in a pET28a expression vector was a gift from Jack Szostak (Harvard Medical School). Mutations to the wild-type PheRS sequence were performed by site directed mutagenesis to generate

PheRS single mutant A294G and double mutant A294G, T251G by Susan Zhizhong Xu. For each point mutation the following DNA oligo pairs (IDT, USA) were used. For PheRS A294G, forward primer: 5'-GAA GTT TAC TCT GGT TTC GGC TTC GGG ATG GGG ATG G-3' and reverse primer: 5'-CCA TCC CCA TCC CGA AGC CGA AAC CAG AGT AAA CTT C-3'.

PheRS A294G, T251G was mutated on the PheRS A294G construct with the forward primer: 5'-CTT CCT ACT TCC CGT TTG GCG AAC CTT CTG CAG AAG TG-3' and reverse primer: 5'-CAC TTC TGC AGA AGG TTC GCC AAA CGG GAA GTA GGA AG-3'. All mutations were verified by DNA sequencing by the Applied Genomics Centre (Department of Medical Genetics, University of Alberta, Canada) or by the Eurofins MWG Operon (Huntsville, AL, USA).

A2.2.3. In vitro Transcription of $tRNA^{Phe}$

See Material and Methods (**2.4**)

A2.2.4. Unnatural Amino Acid Labeling and Click Chemistry on Model Peptide Substrate

The steps to enzymatically aminoacylate $tRNA^{Phe}$ with azido-Phe and subsequently use L/F transferase to transfer the unnatural amino acid to a model polypeptide substrate was performed similarly described by Ebhardt *et al*. except with the following changes (Ebhardt *et al.* 2009). Briefly for a 1 mL reaction, 50 mM Hepes pH 7.5, 50 mM KCl, 15 mM $MgCl_2$, 0.4 mM CTP, 4 mM ATP, 4 mM beta-mercaptoethanol, 0.4 mM

azido-phenylalanine, 8.6 μM $tRNA^{Phe}$, 3.5 μM heavy labeled FREPGLCTWQSLR peptide (standard), 7.1 μM light REPGLCTWQSLR peptide (substrate), 0.43 μM CCA adding enzyme, 0.54 μM PheRS (wild-type, A294G or A294G, T251G) and 8.3 μM L/F transferase were incubated at 37 ºC for 20 hours. Aliquots at various time points were saved and quenched with 10 % acetonitrile and 1 mg/mL BSA in 2 % TFA. The quenched samples were desalted using prepared 3 mL C_{18} cartridge (Agilent Technologies SampliQ C_{18} End-capped). First, 5 volumes of 100 % acetonitrile were added to the cartridges and centrifuged at 2 000 RPM for 10 minutes at 4 ºC. Then the cartridges were equilibrated with 5 volumes of 0.1% TFA. Samples were loaded onto the cartridges, washed with 5 volumes of 0.1 % TFA, and eluted with 3 mL 70 % Acetonitrile with 0.1 % TFA. The samples were completely dried under speed vacuum and resuspended in 10 μL of double distilled water.

For a 100 μL click chemistry reaction, 50 mM Hepes pH 8.0, 100 μM TBTA, 1 mM TCEP, 1 mM $CuSO_4$, 100 μM alkyne (propargylamine, NBD-alkyne (from Dr. Luc Berthiaume University of Alberta), or biotin-alkyne (from Dr. Luc Berthiaume University of Alberta)) and 70 μM azido-Phe-Arg-peptide were incubated at 37 ºC overnight in the dark. The pH of the final mixture is measured to be at 6.5. Then the samples were acetone precipitated by adding 4 volumes of cold acetone and incubate at -20 ºC for more than 1 hour. The samples were centrifuged at 14 000 RPM for 15 min at 4 ºC. The supernatant carefully discarded and the

pellet air dried in the dark. The samples were resuspended in 10 μL of double distilled water.

For MALDI-ToF MS analysis, 1 μL of each aliquot sample was mixed with 1 μL of CHCA solution (saturated CHCA in 50% acetonitrile and 0.2 % TFA), and 1 μL of the mixed solution was spotted onto the MALDI plate. After drying, the samples were analyzed using the Bruker Daltonics Ultraflex (Bruker). Mass spectra were analyzed and exported from the FlexAnalysis software.

A2.2.5. Identification of Potential Biotinylated Proteins in Wild-Type and birA Mutant Lysates

As a preliminary investigation, we performed western blot analysis on wild-type K-12 and *birA* mutant lysates. *birA* mutant strains were purchased from the CGSC *E. coli* Genetics Stock Center (Yale University, New Haven, CT, USA). Cultures of K-12 and the *birA* mutants were grown in 5 mL of LB media overnight at 37 ºC (no antibiotics). The next day, 1 mL of the cells was harvested by centrifugation at 5 000 x g for 8 minutes at 4 ºC. 200 μL of 5 x SDS Loading Buffer were added to the pellet and lysed by micro-sonication (QSonica Q125). Samples were centrifuged at 14 000 RPM for 5 minutes at 4 ºC, heated to 80 ºC for 2 minutes, and loaded to two different 10 % SDS-PAGE gels. After SDS-PAGE gel electrophoresis, one gel was stained in coomassie stain and one gel was transfer to nitrocellulose membrane (LI-COR biosciences).

The nitrocellulose membrane was blocked in 2.5 % fish gelatin in 1 x PBS overnight at 4 ºC, washed with 1 x PBS for three times, and biotinylated proteins were immunoblotted with IRDye 680®-streptavidin (LI-COR) for 45 minutes at room temperature. The gel and nitrocellulose membrane were scanned (LI-COR Odyssey® Infrared Imaging System).

A2.2.6. K-12 and AB313-136 Fractionation

K-12 and AB313-136 cells were grown in a 50 mL LB media overnight at 37 ºC. The cells were transferred to 500 mL of LB media and grown at 37 ºC until stationary phase ($O.D._{600nm}$ > 1.0). The cells were harvested by centrifugation and washed in 10 mM Tris (pH 7). The cells were resuspended in 20 mL Lysis Buffer A (see **2.3** – 1 x HisTrap buffer). Cells were lysed by adding 0.2 g of lysosome and sonicated (QSonica Q125). The lysed cells were centrifuged at 6000 x g for 15 minutes at 4 ºC. Carefully decant the supernatant (supernatant 1) into falcon tube and placed on ice. The pellet was resuspended in 20 mL Lysis Buffer A and centrifuged again at 6 000 x g for 15 minutes at 4 ºC. The resulting pellet (pellet 2) was resuspended in Resuspension Buffer (50 mM Tris-HCl (pH 7.4), 800 mM NaCl, 20 % glycerol) stirring at 4 ºC and labeled as 'debris fraction'. Meanwhile the resulting supernatant (supernatant 2) was combined with supernatant 1, and centrifuged at 40 000 RPM for 1.5 hours at 4 ºC.

The resulting supernatant (supernatant 3) was collected and labeled as 'cytosolic fraction'. And the resulting pellet (pellet 3) were

resuspended in 25 mL of Resuspension Buffer with 1 % *n*-dodecyl-β-D-maltoside (β-D-DDM) and stirred for 45 minutes at 4 ºC. Then the sample was centrifuged at 40 000 RPM for 45 minutes at 4 ºC. The resulting supernatant (supernatant 4) was labeled as 'peripheral fraction'. Meanwhile the resulting pellet (pellet 4) was resuspended in 1 mL of Solubilization Buffer and labeled as 'integral fraction'. 5 x SDS Loading Buffer were added to the cytosolic, peripheral, integral, and debris fractions. SDS-PAGE and immunoblot were performed similarly as described above.

A2.2.7. Biotinylated Proteins Immnoprecipitation

We first tested the CaptAvidin (Life Technologies), NeutrAvidin (Pierce Net), and streptavidin (Sigma) agarose beads according to manufacturer's instructions. We found that streptavidin agarose beads give the most reproducible. To identify potential background proteins that bind to streptavidin beads, we have performed preliminary streptavidin immunoprecipitation experiments with wild-type K-12 lysates. Briefly, 1 mL of streptavidin agarose beads are washed twice with 5 mL of 1 x PBS and centrifuge at 8 000 RPM for 10 minutes at 4 ºC. 10 mL of cleared lysates (an aliquot taken as 'input') were incubated with the streptavidin beads for 15 minutes at room temperature. Centrifuge and save supernatant as 'flow through'. Wash the beads four times with 1 x PBS for 15 minutes and save supernatant as 'wash'. Elute with 2.5 mL of 5 mM biotin in 1 x PBS for 2 hours at room temperature. Centrifuge and save

supernatant as ‘elution’. SDS-PAGE and immunoblot were performed similarly as described above.

Elution and wash lanes were cut, reduced, alkylated, and digested with trypsin as described in **A1.2.5**. The samples then are subjected to in-gel LC-MS/MS protein identification by the Thermo LTQ orbitrap XL (Thermo Scientific) and searched against the in-house SEQUEST server. DAVID (http://david.abcc.ncifcrf.gov/) and STRING (http://string-db.org/) online analyses were used to identify pathways and interaction proteins, respectively.

A2.2.8. ∆aat Deletion Strain Confirmation

The ∆*aat* strain (ASAP ID: ABE-0003009, aat::Tn5 (KAN-2) at position 103 in (+) orientation) was purchased from ASAP of University of Wisconsin. To confirm the Tn5 insertion that generates the ∆*aat* strain, whole cell PCR was performed from wild-type K-12 cells and ∆*aat* cells using forward sequencing primer 5'-CTC ACG CAG AAC TGC TTG CCA GAC-3' and reverse sequencing primer 5'-CCT GAC GCG AAA AAT TCC TTA TCG G-3'. Amplified gene fragments were analyzed by 0.8% agarose gel electrophoresis, gel extracted (QIAGEN) and purified, then confirmed by sequencing (Department of Medical Genetics, University of Alberta).

A2.2.9. ∆aat Growth Phenotype during Stress

To test the growth phenotype between K-12 and Δ*aat* cells, we tested the cells under temperature and nutritional stress in liquid and solid media. Briefly, 2 mL overnight K-12 and Δ*aat* cultures grown in LB media or M9 minimal media (Atlas 1993) were inoculated to pre-warmed (31, 37, 42, and 45 ºC) 125 mL of the same media. An aliquot of 1 mL is taken and measured the optical density at 600 nm every half hour. Graphs are plotted with $O.D._{600nm}$ against time. Solid LB agar and M9 minimal agar were also used to test the growth phenotype between K-12 and *Δaat*. 5 μL of 10-fold serial dilutions (10^1 to 10^7 dilution) of an 2 mL overnight culture were spotted on the LB agar and M9 minimal agar plates and grown overnight at 37 ºC.

A2.3. Results

A2.3.1. Unnatural Amino Acid Labeling and Click Chemistry on Model Peptide Substrate

For the unnatural amino acid labeling procedure, our laboratory has already published the steps to enzymatically aminoacylate $tRNA^{Phe}$ with azido-Phe and subsequently use L/F transferase to transfer the unnatural amino acid to a model polypeptide substrate (**Figure A2-2**) (Ebhardt *et al.* 2009). Since then, I have further optimized the yield of azido-Phe-Arg-peptide by changing the concentrations of each component. Each stage of the reaction and the final products has been confirmed using MALDI-ToF MS. I have found that PheRS single mutant A295G aminoacylates

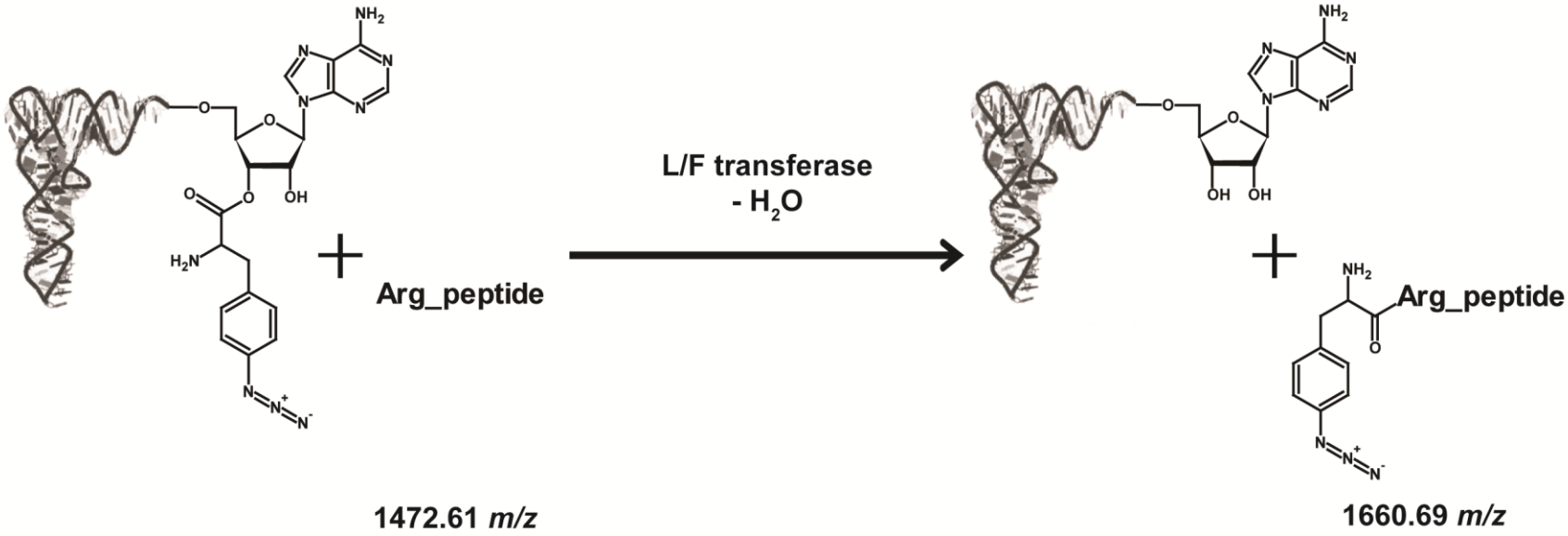

Figure A2-2: A schematic of the azido-Phe addition reaction to model Arg-peptide catalyzed by L/F transferase.

azido-Phe onto $tRNA^{Phe}$ more quantitatively than the previously reported PheRS double mutant A295G, T251G. **Figure A2-3** shows the superimposed mass spectra of an azido-Phe addition reaction after 2 min (red) and 20 hours (blue). The expected product azido-Phe-Arg-peptide mass peak is observed at 1661.01 *m/z*, which confirms the successful addition of azido-Phe onto the N-terminus of the model substrate Arg-peptide (Arg-peptide 1472.61 *m/z* + azido-Phe 206.10 *m/z* – H_2O 18.02 *m/z* = azido-Phe-Arg-peptide 1660.69 *m/z* [$M+H^+$] expected). Since the azido group is sensitive to ultra-violet (UV) light and a UV laser is used during MALDI analysis, side products were observed with one or two nitrogen atom loss.

Here I demonstrate that the azido-Phe modified polypeptide can be derivatized using the click chemistry procedure. **Figure A2-4** shows a schematic of click chemistry between azido-Phe-Arg-peptide and biotin-alkyne probe. **Figure A2-5** shows the superimposed mass spectra of the click chemistry reaction between azido-Phe-Arg-peptide and three different alkyne probes tested: propargylamine (blue), nitrobenzoxadiazole (NBD)-alkyne (red), and biotin-alkyne (green). The expected products peptide mass peak were observed at 1715.92, 1879.18, 2188.21 *m/z*, which confirms successful click chemistry reactions.

A2.3.2. Additional Biotinylated Proteins Enriched in Stationary Phase and Cytoplasmic Fraction of E. coli

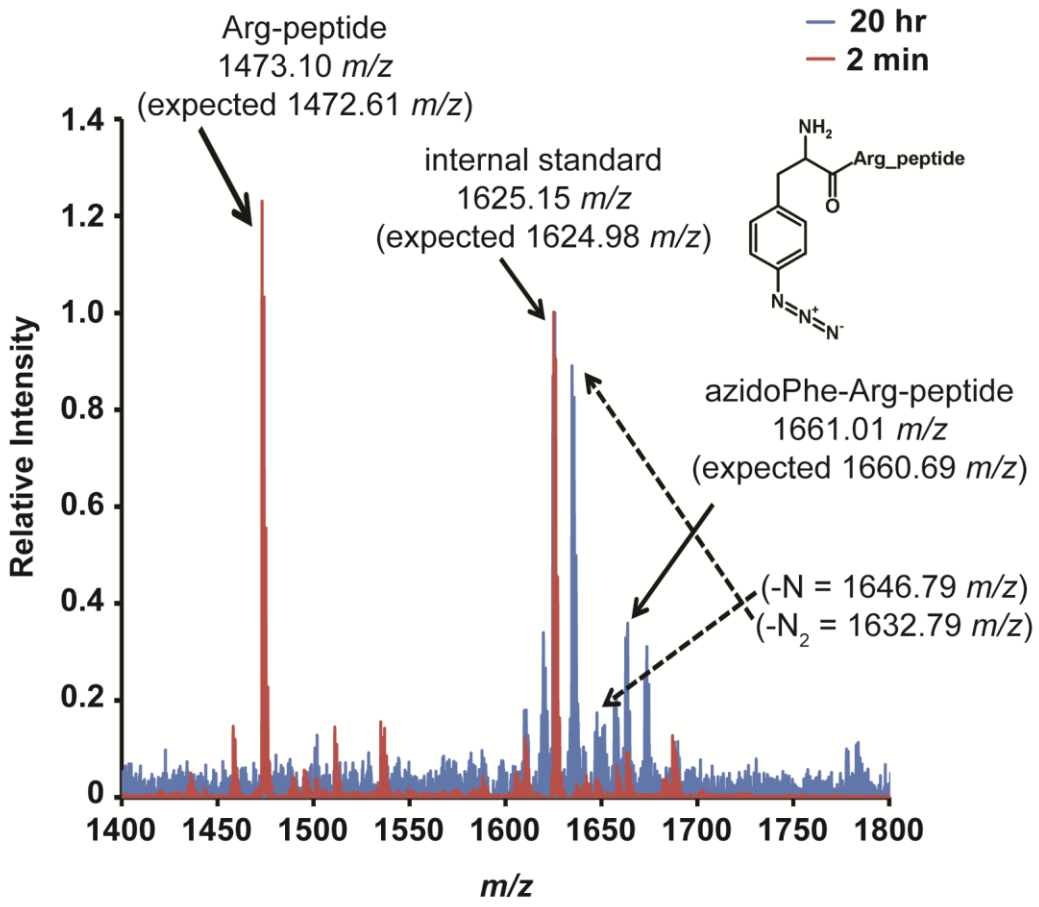

Figure A2-3: Superimposed mass spectra of azido-Phe addition to Arg-peptide after 2 minutes (red) and 20 hours (blue). The azidoPhe-Arg-peptide product is observed at 1661.01 *m/z* (Arg-peptide 1472.61 *m/z* + azido-Phe 206.10 *m/z* – H_2O 18.02 *m/z* = azido-Phe-Arg-peptide 1660.69 *m/z* [$M+H^+$] expected).

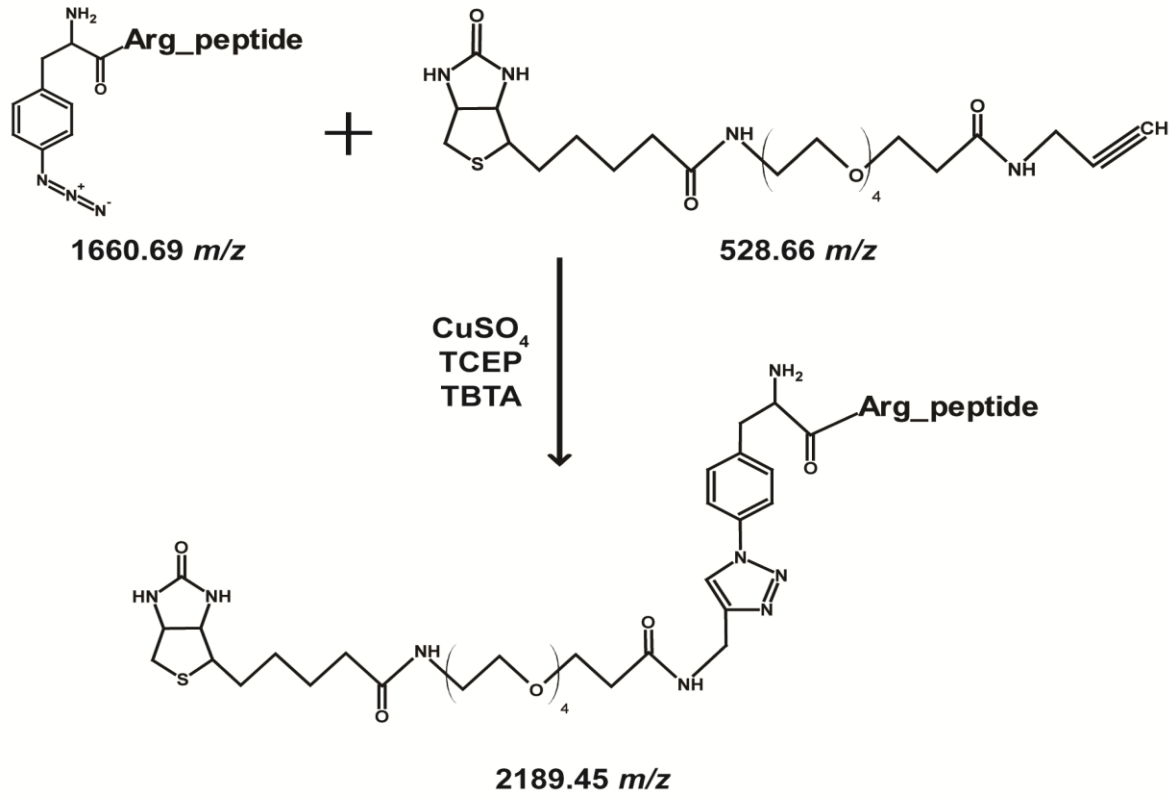

Figure A2-4: A schematic of the click chemistry reaction between azido-Phe-Arg-peptide and biotin-alkyne.

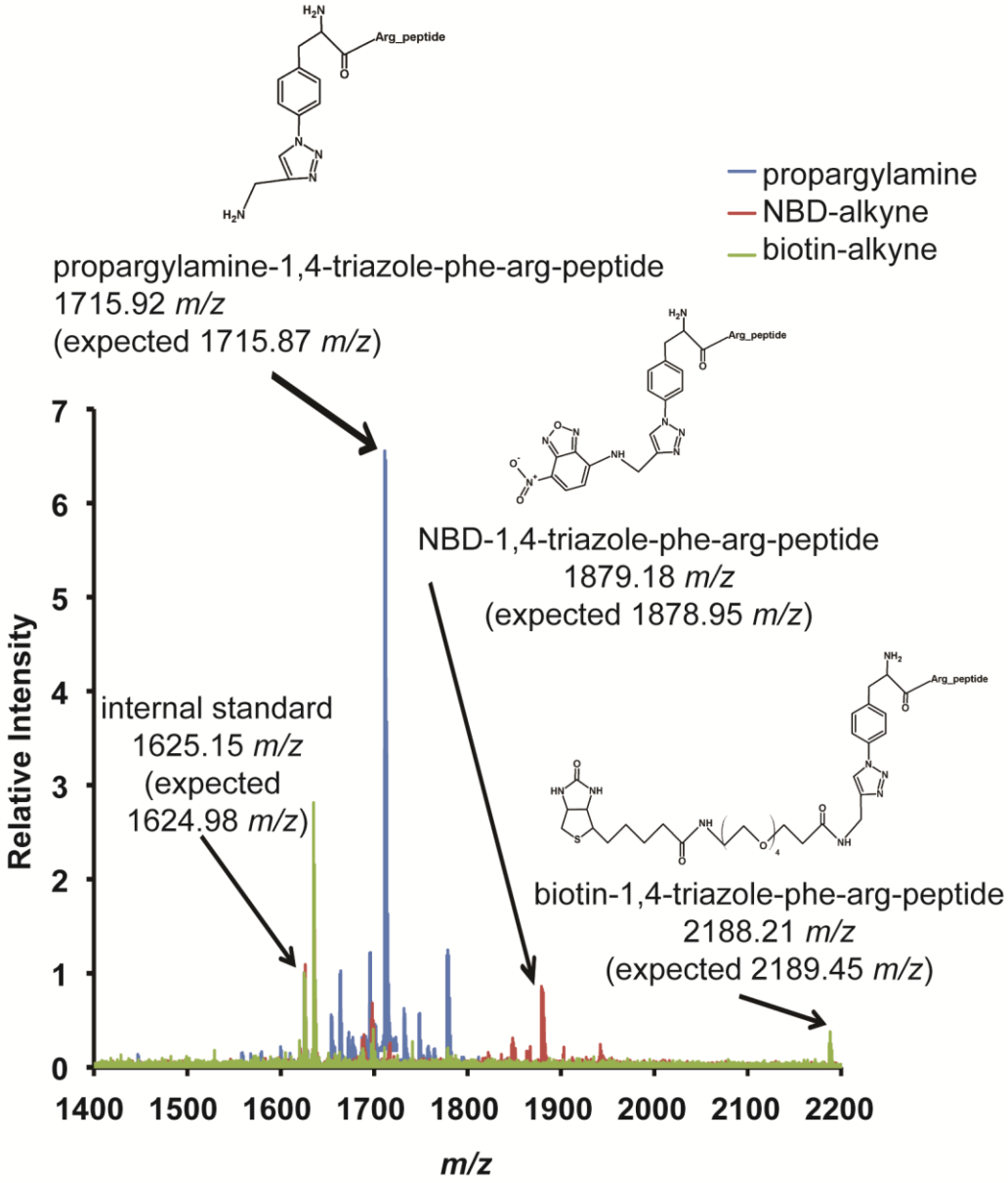

Figure A2-5: Superimposed mass spectra of click chemistry reactions of azido-Phe-Arg-peptide with three different alkynes: propargylamine (blue), NBD-alkyne (red), and biotin-alkyne (green). Propargylamine-1,4-triazole-phe-arg-peptide product is observed at 1715.92 m/z, NBD-alkyne-1,4-triazole-phe-arg-peptide product is observed at 1879.18 m/z, and biotin-alkyne-1,4-triazole-phe-arg-peptide product is observed at 2188.21 m/z.

The isolation procedure depends heavily on biotinylation. There is a single known biotinylated protein in *E. coli*, the biotin carboxyl carrier protein (BCCP) subunit of acetyl-CoA carboxylase (~17kDa) (Chapman-Smith and Cronan 1999), but the presence of others could result in false positive identifications. We have evaluated this and have demonstrated that there are a variety of low abundant biotinylated proteins in *E. coli*.

As a preliminary investigation, we performed western blot analysis on wild-type K-12 cytoplasmic cleared lysates from different growth phases. The cleared cytoplasmic lysates of the different cultures were first separated via SDS-PAGE followed by immunoblotting with $IRDye_{680}^{®}$-streptavidin. Biotin-Avidin is an extremely strong non-covalent interaction ($K_a = 10^{15}\ M^{-1}$). We observed that there are more additional bands immunoblot by $IRDye_{680}^{®}$-streptavidin in the stationary phase than other growth phases (data not shown). Subsequent experiments were done on stationary phase lysates.

The gene *birA* encodes the bifunctional protein BirA, which acts both as a biotin-operon repressor and synthesizes the co-repressor, acetyl-CoA: carbon dioxide ligase. BirA also activates biotin to form biotinyl-5'-adenylate and transfers the biotin moiety to biotin-accepting proteins. *birA* mutants (strain AB313-136 (Eisenberg 1975) and strains BM4056, BM4062, BM4064, BM4092 (Barker and Campbell 1980)) have modified BirA activity and are useful as controls in the identification of other biotinylated proteins.

Figure A2-6 shows the immunoblot and SDS-PAGE of *E. coli* wild-type K-12 and *birA* mutant cleared, stationary-phase, cytoplasmic lysates. The immunoblot shows additional bands in the higher molecular weights of the K-12 lysates suggesting the presence of other biotinylated proteins. The band at ~20 kDa is identified to be BCCP of acetyl-CoA carboxylase by in-gel LC-MS/MS. Of the five *birA* mutant strains, we observed that strain AB313-136 has enhanced biotinylation while strain BM4056 (birA85) has significantly diminished natural biotinylation. These two strains may be good candidates for positive and negative controls.

To systematically identify the localization of other biotinylated proteins, we have performed an alternative fractionation procedure on the stationary-phase lysates. **Figure A2-7** shows the experimental flowchart on the *E. coli* protein fractionation procedure to separate proteins into debris, cytosolic, peripheral membrane, and integral membrane fractions. Followed by SDS-PAGE and immunoblot analysis on the fractionated samples, we observed that most of the additional biotinylated proteins localize in the cytoplasm (**Figure A2-8**). Subsequent procedures focus on stationary phase lysates of the cytoplasmic fraction alone.

A2.3.3. Identification of Background Biotinylated Proteins via Streptavidin Agarose Immunoprecipitation

To identify potential biotinylated proteins in the *E. coli* stationary-phase cytoplasmic fraction, we performed an immunoprecipitation procedure prior to identifying other biotinylated proteins. We have tested

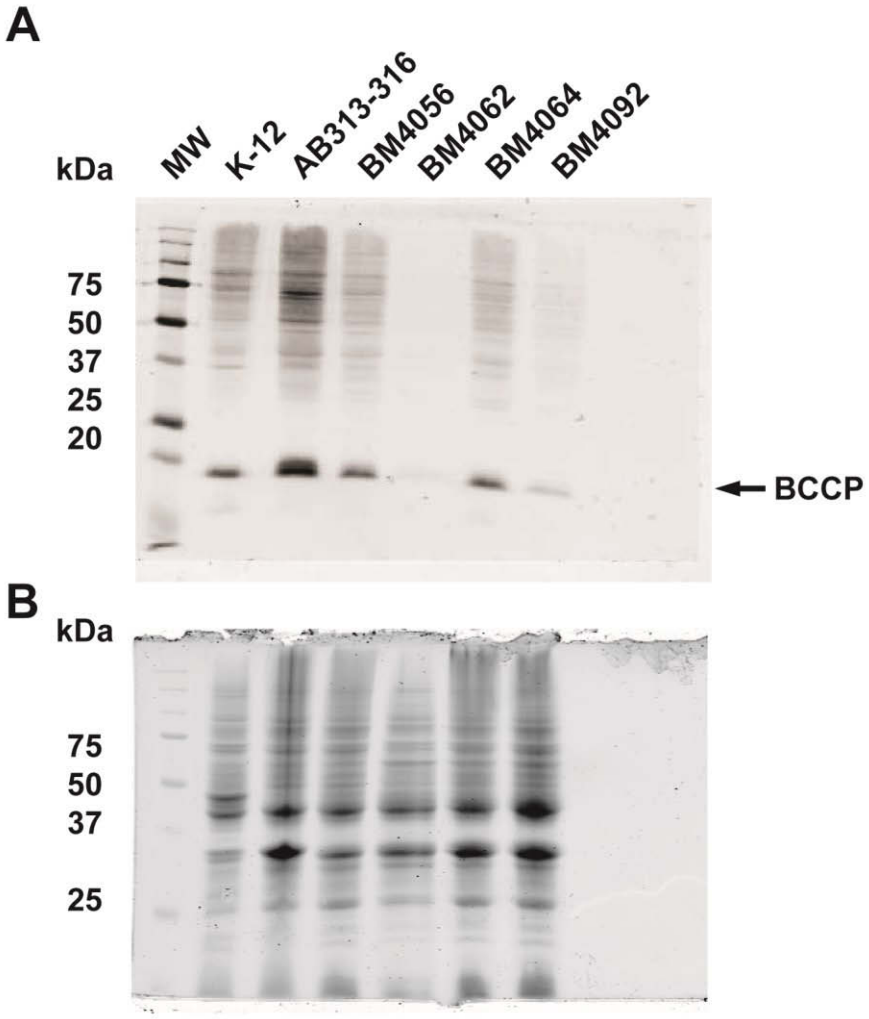

Figure A2-6: A) Immunoblot with IRDye 680®-streptavidin and B) SDS-PAGE of wild type K-12 and birA mutant strains (AB313-316, BM4056, BM4062, BM4064, and BM4092) stationary phase lysates.

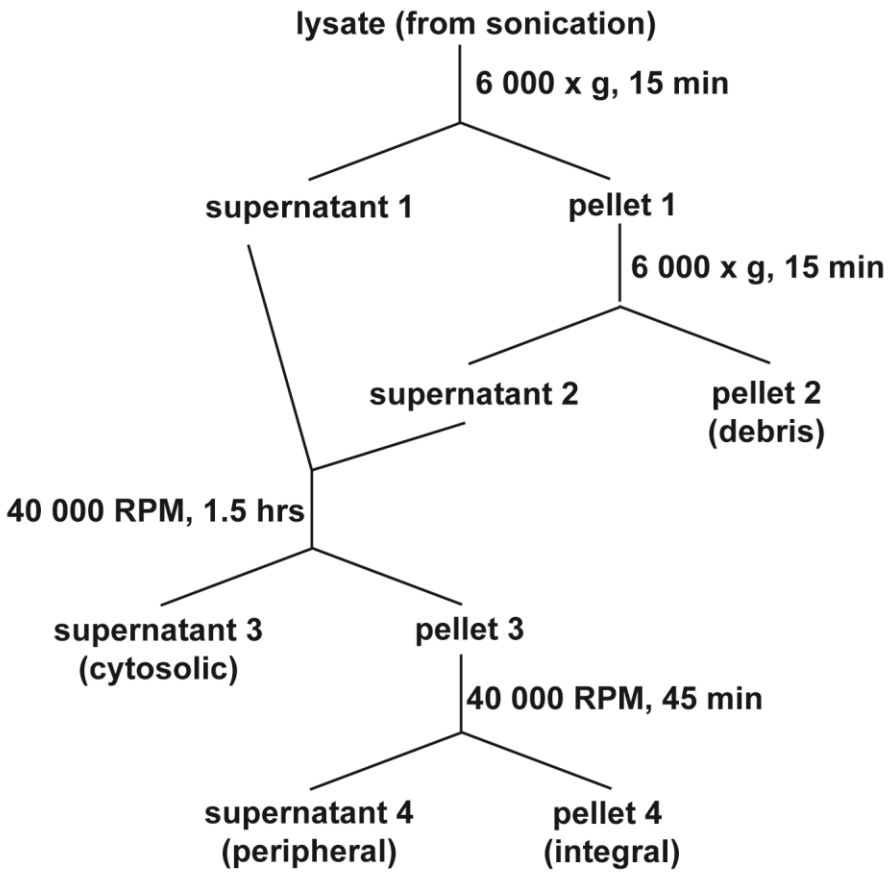

Figure A2-7: Schematic of the lysate fractionation protocol.

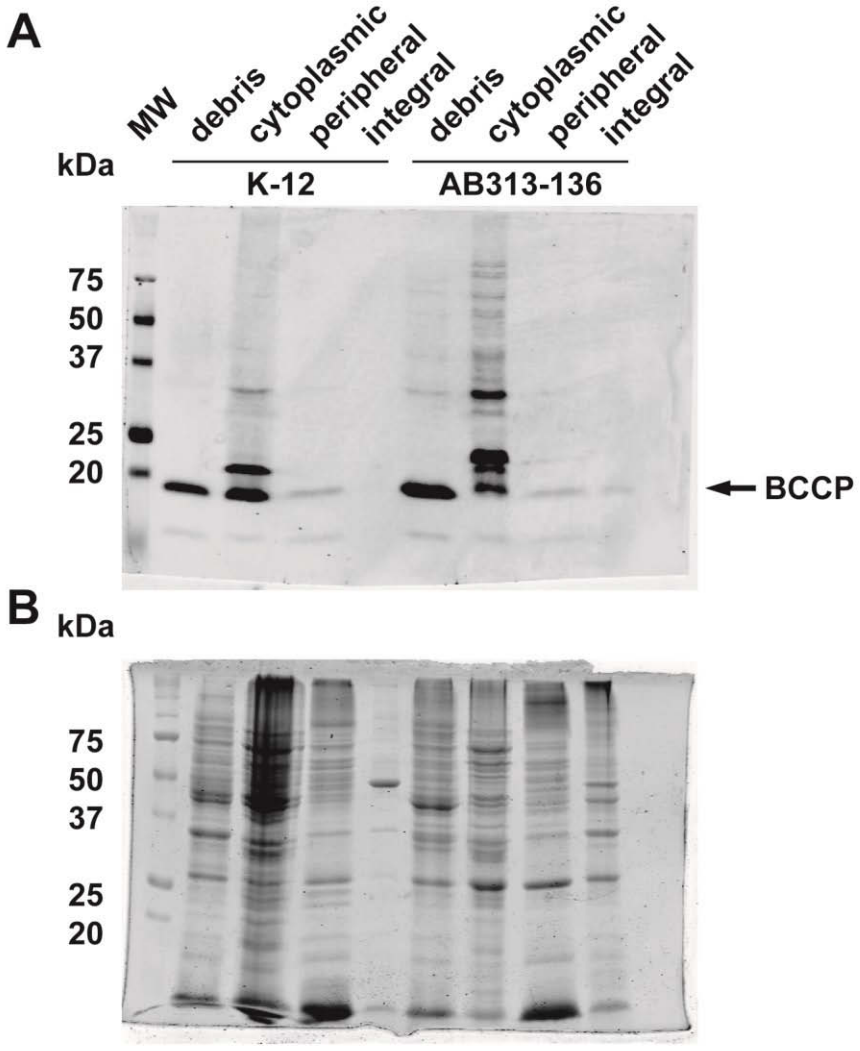

Figure A2-8: A) Immunoblot with IRDye 680®-streptavidin and B) SDS-PAGE of wild-type K-12 and *birA* mutant strain AB313-136 following lysate fractionation. Additional biotinylated proteins are observed in the cytosolic fractions.

CaptAvidin, NeutrAvidin, and streptavidin agarose beads and found that streptavidin agarose beads provide the most specific and reproducible elution profiles (data not shown). We have then optimized the binding and elution buffer compositions. We observed that the manufacturer's suggested buffers containing 2% SDS and 0.4 M urea are too harsh of a denaturant, and elute proteins non-specifically. After testing several buffers, we found that 5 mM biotin in 1 x PBS is a suitable elution buffer that elutes biotinylated proteins via competition but remains gentle.

Figure A2-9 shows the immunoblot and SDS-PAGE of the streptavidin immunoprecipitation fractions. As expected, we observed additional protein bands on the immunoblot in the elution fraction. Although the wash fractions show little biotinylated proteins on the immunoblot, but the presence of proteins observed in the SDS-PAGE suggests that the wash fraction should also be subject to protein identification as a negative control. Thus, whole lanes of wash 1 and elution were excised, reduced and alkylate, and subject to trypsin digestion (see procedure in **Figure A1-6**).

Upon subtracting proteins identified from the washes, we found 238 proteins (344 – 106 = 238) eluted in the first trial and 6 proteins (50 - 44 = 6) eluted in the second trial. The proteins identified are listed in **Table A2-1**. The low amounts of protein eluted from the beads on the second trial may be explained by the weak lysates during preparation. Despite the low amounts detected, there are 4 proteins in common to both replicates and

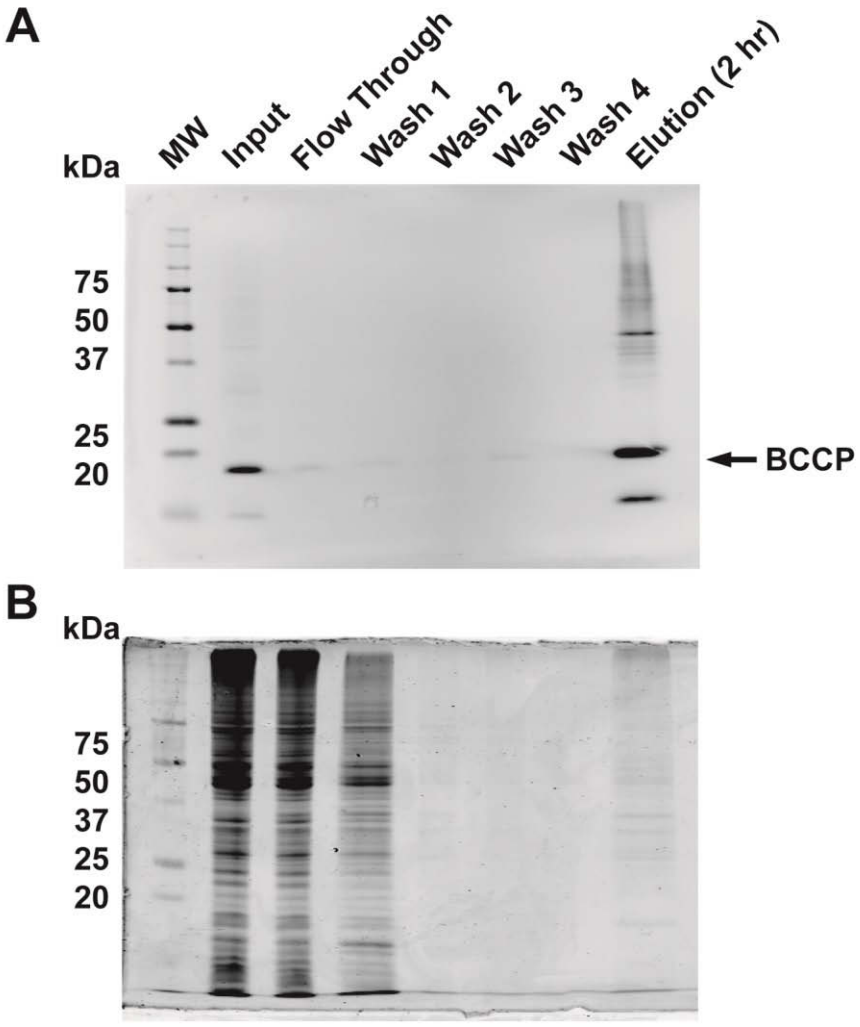

Figure A2-9: A) Immunoblot and B) SDS-PAGE of streptavidin immunoprecipitation.

Table A2-1: List of proteins identified by streptavidin immunoprecipitation (wash controls subtracted).

Accession Number	Protein	# peptide	% Coverage	SEQUEST score
Trial 1				
P0A9A6	Cell division protein FtsZ	20	74	296
C4ZXE9	Glycine--tRNA ligase beta subunit	25	48	249
P69776	Major outer membrane lipoprotein Lpp	3	49	190
P0ACJ8	Catabolite gene activator	11	44	186
P0A9X4	Rod shape-determining protein MreB	10	42	160
C4ZXL1	ADP-L-glycero-D-manno-heptose-6-epimerase	10	35	144
C4ZYJ2	Elongation factor 4	13	23	126
C4ZSW3	N-acetylneuraminate lyase	7	33	121
P0AES6	DNA gyrase subunit B	14	25	110
P0ABH9	ATP-dependent Clp protease ATP-binding subunit ClpA	16	25	106
C4ZSQ9	Translation initiation factor IF-2	14	20	105
P09152	Respiratory nitrate reductase 1 alpha chain	24	30	102
C4ZYU4	Protein RecA	9	37	96
C4ZX73	Uracil phosphoribosyltransferase	6	38	96
P69407	Capsular synthesis regulator component B	9	51	93
P00960	Glycine--tRNA ligase alpha subunit	5	21	92
C4ZQW9	Bifunctional protein HldE	10	32	87
P0ABD8	**Biotin carboxyl carrier protein of acetyl-CoA carboxylase**	**3**	**24**	**86**
C5A0R1	Soluble pyridine nucleotide transhydrogenase	11	39	79
P69797	PTS system mannose-specific EIIAB component	12	56	79
P07014	Succinate dehydrogenase iron-sulfur subunit	7	40	78
C4ZTN0	D-amino acid dehydrogenase small subunit	9	30	73
P37440	Oxidoreductase UcpA	8	42	70
C4ZZ11	ATP synthase gamma chain	10	38	70
P0A912	Peptidoglycan-associated lipoprotein	4	36	67

P37903	Universal stress protein F	3	25	67
P0AAC0	Universal stress protein E	7	31	63
P0A9Q1	Aerobic respiration control protein ArcA	8	37	63
C4ZQC8	Curved DNA-binding protein	8	35	61
P0ADY1	Peptidyl-prolyl cis-trans isomerase D	15	33	59
C4ZY85	Adenosine deaminase	5	26	57
P39177	Universal stress protein G	4	41	55
P09546	Bifunctional protein PutA	14	17	54
P69908	Glutamate decarboxylase alpha	8	22	51
P10121	Signal recognition particle receptor FtsY	12	29	51
C5A0L1	S-adenosylmethionine synthase	8	29	50
C4ZUG8	50S ribosomal protein L16	3	32	50
P0AA16	Transcriptional regulatory protein OmpR	6	33	48
P37636	Multidrug resistance protein MdtE	12	37	46
C4ZUG5	**50S ribosomal protein L14**	**5**	**38**	**44**
P0ACY3	Uncharacterized protein YeaG	11	21	44
C4ZRS7	Acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha	5	26	43
C4ZX90	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase	11	38	42
P0AE06	**Acriflavine resistance protein A**	**10**	**36**	**41**
P0AG20	GTP pyrophosphokinase	4	7	41
P0AEP3	UTP--glucose-1-phosphate uridylyltransferase	5	26	41
C4ZU83	Anaerobic glycerol-3-phosphate dehydrogenase subunit B	6	22	41
C5A130	Maltoporin	8	30	39
P33012	DNA gyrase inhibitor	3	29	38
P28635	D-methionine-binding lipoprotein MetQ	5	31	38
P0A905	Outer membrane lipoprotein SlyB	4	41	37
P37744	Glucose-1-phosphate thymidylyltransferase 1	5	22	37
P16456	Selenide, water dikinase	6	22	37
P0AFG0	Transcription antitermination protein NusG	6	43	37

C4ZUG3	50S ribosomal protein L5	3	22	37
P00363	Fumarate reductase flavoprotein subunit	6	14	36
P37194	Outer membrane protein slp	3	20	34
C4ZRJ3	Protein translocase subunit SecA	8	10	33
C4ZQ11	Leucyl/phenylalanyl-tRNA--protein transferase	4	18	31
P21513	Ribonuclease E	8	10	31
P07017	Methyl-accepting chemotaxis protein II	9	22	30
P0ACB0	Replicative DNA helicase	9	26	30
P03023	Lactose operon repressor	4	12	30
P0ACE0	Hydrogenase-2 large chain	6	18	30
P66948	TPR repeat-containing protein YfgC	6	20	29
P0AGD7	Signal recognition particle protein	7	25	29
P11349	Respiratory nitrate reductase 1 beta chain	8	22	28
P76372	Chain length determinant protein	7	32	28
P0ADK0	Uncharacterized protein YiaF	5	35	28
P68187	Maltose/maltodextrin import ATP-binding protein MalK	7	25	28
P76387	Tyrosine-protein kinase wzc	2	4	28
P42630	L-serine dehydratase TdcG	4	14	27
P0ABU2	GTP-dependent nucleic acid-binding protein EngD	5	17	27
C4ZRP7	Glutamate-1-semialdehyde 2,1-aminomutase	8	29	26
P30958	Transcription-repair-coupling factor	10	11	26
P0AAB6	UTP--glucose-1-phosphate uridylyltransferase	3	15	26
P0ABC3	Modulator of FtsH protease HflC	5	18	25
P0AAI3	ATP-dependent zinc metalloprotease FtsH	6	14	25
P0AGF6	L-threonine dehydratase catabolic TdcB	5	17	25
P0A996	Anaerobic glycerol-3-phosphate dehydrogenase subunit C	6	17	24
C4ZYY3	Membrane protein insertase YidC	5	17	24
C4ZQF4	tRNA (cmo5U34)-methyltransferase	7	40	24
P0AAG8	Galactose/methyl galactoside import ATP-	8	19	24

	binding protein MglA			
B1XBE0	Chaperone protein DnaJ	4	14	23
C4ZX86	GTPase Der	6	14	23
P08506	D-alanyl-D-alanine carboxypeptidase DacC	6	25	23
B1X8Z8	NADH-quinone oxidoreductase subunit C/D	7	14	23
P08194	Glycerol-3-phosphate transporter	2	4	23
P09127	Putative uroporphyrinogen-III C-methyltransferase	8	32	23
P02942	Methyl-accepting chemotaxis protein I	7	20	22
C4ZXV1	UvrABC system protein B	2	4	22
P0ABZ6	Chaperone SurA	5	18	22
P0AG90	Protein translocase subunit SecD	8	16	21
P00550	PTS system mannitol-specific EIICBA component	5	10	21
P23836	Transcriptional regulatory protein PhoP	6	37	21
P09323	PTS system N-acetylglucosamine-specific EIICBA component	2	5	21
P76193	Probable L,D-transpeptidase YnhG	5	29	21
P37665	Inner membrane lipoprotein YiaD	3	26	21
P31802	Nitrate/nitrite response regulator protein NarP	5	32	21
P77804	Protein YdgA	7	17	21
P0A698	UvrABC system protein A	7	10	20
P69924	Ribonucleoside-diphosphate reductase 1 subunit beta	4	15	20
P03841	Maltose operon periplasmic protein	4	24	20
B1X927	Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta	6	22	20
P69828	Galactitol-specific phosphotransferase enzyme IIA component	4	35	19
P00452	Ribonucleoside-diphosphate reductase 1 subunit alpha	4	6	19
C4ZZN7	DNA mismatch repair protein MutS	6	11	19
P16095	L-serine dehydratase 1	4	10	19
P0ABH0	Cell division protein FtsA	4	19	19
P37637	Multidrug resistance protein MdtF	7	9	19

P08192	Bifunctional protein FolC	5	19	19
C4ZQ50	Chromosome partition protein MukB	8	6	19
Q46802	Uncharacterized sigma-54-dependent transcriptional regulator YgeV	4	9	18
P0AEE1	Protein DcrB	3	22	18
P19932	Hydrogenase-1 operon protein HyaF	3	23	18
C4ZT99	Glutamate 5-kinase	4	17	18
P11557	Protein DamX	5	15	18
P0ADW3	Putative cytochrome d ubiquinol oxidase subunit 3	5	52	18
P0AFR4	Uncharacterized protein YciO	4	24	17
P32176	Formate dehydrogenase-O major subunit	4	8	17
P0AF08	Protein mrp	3	11	17
C4ZS48	UPF0227 protein YcfP	4	31	17
P0AFI2	DNA topoisomerase 4 subunit A	5	8	16
P0A959	Glutamate-pyruvate aminotransferase AlaA	3	9	16
P46837	Protein YhgF	4	6	15
P77735	Uncharacterized oxidoreductase YajO	3	13	15
C4ZZ22	Aspartate-ammonia ligase	2	8	15
P77541	Methylisocitrate lyase	4	16	15
P0C0V0	Periplasmic serine endoprotease DegP	4	13	15
P52108	Transcriptional regulatory protein RstA	4	18	15
P0ADZ7	UPF0092 membrane protein YajC	4	27	14
P20083	DNA topoisomerase 4 subunit B	3	5	14
C4ZSR2	Argininosuccinate synthase	6	16	14
B1X9W4	ATP synthase subunit b	5	40	14
P77737	Oligopeptide transport ATP-binding protein OppF	4	13	14
P0ABJ1	Ubiquinol oxidase subunit 2	3	16	14
C4ZWP8	NAD-dependent malic enzyme	4	9	14
C4ZX51	Phosphoribosylaminoimidazole-succinocarboxamide synthase	4	21	13
B1X658	Alpha,alpha-trehalose-phosphate synthase [UDP-forming]	6	12	13
P0AED0	Universal stress protein A	2	23	12

P0A903	Outer membrane protein assembly factor BamC	3	13	12
P69805	Mannose permease IID component	2	8	12
P31224	Acriflavine resistance protein B	4	5	12
P0A9H7	Cyclopropane-fatty-acyl-phospholipid synthase	2	7	12
C4ZYN1	Protein GrpE	2	14	12
P0A917	Outer membrane protein X	2	20	11
P07363	Chemotaxis protein CheA	3	6	11
C4ZSQ8	**Ribosome-binding factor A**	**3**	**21**	**11**
P64604	Probable phospholipid ABC transporter-binding protein MlaD	2	11	10
P0AC47	Fumarate reductase iron-sulfur subunit	3	12	10
P26459	Cytochrome bd-II oxidase subunit 1	3	9	10
C5A134	Glycerol-3-phosphate acyltransferase	4	8	10
P07003	Pyruvate dehydrogenase [ubiquinone]	2	4	10
P13445	RNA polymerase sigma factor RpoS	4	17	10
P0AE88	Transcriptional regulatory protein CpxR	3	20	10
C5A020	Fatty acid oxidation complex subunit alpha	3	6	9
P0A6E9	ATP-dependent dethiobiotin synthetase BioD 2	3	16	9
P0AEB2	D-alanyl-D-alanine carboxypeptidase DacA	3	8	9
P37648	Protein YhjJ	3	9	9
P0AAA1	Inner membrane protein YagU	3	21	9
C4ZRR9	Outer membrane protein assembly factor BamA	3	5	9
P00579	RNA polymerase sigma factor RpoD	3	5	9
C4ZUF2	30S ribosomal protein S11	3	28	8
C4ZUH0	50S ribosomal protein L22	3	34	8
P39342	Uncharacterized protein YjgR	3	8	8
P0A964	Chemotaxis protein CheW	2	17	8
C4ZZE0	UPF0229 protein YeaH	2	5	8
P0AE01	tRNA (cytidine/uridine-2'-O-)-methyltransferase TrmJ	3	16	8
P31554	LPS-assembly protein LptD	3	6	8
P27303	Multidrug resistance protein A	3	10	8

C4ZUH2	50S ribosomal protein L2	2	12	7
P26646	Putative quinone oxidoreductase YhdH	3	16	7
P69786	PTS system glucose-specific EIICB component	2	5	7
P10384	Long-chain fatty acid transport protein	3	10	7
P0AF28	Nitrate/nitrite response regulator protein NarL	2	12	7
P0AAX8	Probable L,D-transpeptidase YbiS	2	12	7
P76594	Uncharacterized protein YfiQ	3	5	7
P0ADA5	Uncharacterized lipoprotein YajG	2	10	7
C4ZYY4	tRNA modification GTPase MnmE	3	8	7
P77499	Probable ATP-dependent transporter SufC	3	15	6
P0AEU7	Chaperone protein skp	2	17	6
P10346	Glutamine transport ATP-binding protein GlnQ	2	12	6
P0AG27	Uncharacterized protein YibN	2	17	6
P0A9K3	PhoH-like protein	2	12	6
P0AC23	Probable formate transporter 1	3	12	6
C4ZYH8	50S ribosomal protein L20	2	16	6
P0AAB8	Universal stress protein D	2	18	6
P77717	Uncharacterized lipoprotein YbaY	2	22	6
P06709	Bifunctional protein BirA	2	12	6
C4ZQY5	G/U mismatch-specific DNA glycosylase	2	21	6
P21367	Uncharacterized protein YcaC	2	11	6
C4ZTH2	6,7-dimethyl-8-ribityllumazine synthase	2	21	6
P06710	DNA polymerase III subunit tau	2	4	6
C4ZRI6	UDP-N-acetylmuramate--L-alanine ligase	2	4	6
C4ZZ48	ATP-dependent RNA helicase RhlB	2	5	6
P00934	Threonine synthase	2	7	6
C4ZZI6	Protease HtpX	2	12	5
C4ZYG5	Phosphoenolpyruvate synthase regulatory protein	2	7	5
C4ZQ48	Chromosome partition protein MukF	2	5	5
P0A9M0	Lon protease	3	4	5

P32131	Oxygen-independent coproporphyrinogen-III oxidase	2	5	5
P76010	Flagellar brake protein YcgR	2	9	5
P0ACB7	Protein HemY	2	6	5
P75818	Uncharacterized lipoprotein YbjP	2	12	5
P00393	NADH dehydrogenase	2	5	5
C4ZQE4	Holliday junction ATP-dependent DNA helicase RuvB	2	9	5
P0A9E5	Fumarate and nitrate reduction regulatory protein	2	8	5
P39396	Inner membrane protein YjiY	2	4	5
P0ACP1	Catabolite repressor/activator	2	9	5
P37749	Beta-1,6-galactofuranosyltransferase WbbI	2	7	5
P02916	Maltose transport system permease protein MalF	2	4	5
C4ZVU3	Cell division protein ZipA homolog	2	10	5
C4ZR48	tRNA dimethylallyltransferase	2	11	5
P0C0S1	Small-conductance mechanosensitive channel	2	8	5
P0AA53	Protein QmcA	2	9	5
C4ZUD0	NADH-quinone oxidoreductase subunit B	2	10	5
P76576	UPF0070 protein YfgM	2	21	5
P0AG00	Lipopolysaccharide biosynthesis protein WzzE	2	8	5
P77774	Outer membrane protein assembly factor BamB	2	8	5
C4ZXA1	Chaperone protein HscA	2	4	5
P0AGD3	Superoxide dismutase [Fe]	2	12	5
P76578	Uncharacterized lipoprotein YfhM	2	1	4
P69739	Hydrogenase-1 small chain	2	8	4
C4ZTM8	Fatty acid metabolism regulator protein	2	12	4
C4ZW75	C4-dicarboxylate transport protein	2	6	4
C4ZRB8	Aspartate carbamoyltransferase	2	6	4
C4ZYI9	GTPase Era	2	8	4
P77739	Uncharacterized protein YniA	2	12	4
C4ZYJ0	Ribonuclease 3	2	8	4

C4ZWB5	Lipoyl synthase	2	10	4
B1XEA0	Transcriptional regulator LsrR	2	7	4
C4ZRJ6	Cell division protein ZapD	2	8	4
P29131	Cell division protein FtsN	2	6	4
P04983	Ribose import ATP-binding protein RbsA	2	4	4
P39099	Periplasmic pH-dependent serine endoprotease DegQ	2	5	4
P0AFK0	Protein PmbA	2	5	4
P0AFX9	Sigma-E factor regulatory protein RseB	2	8	4
P24554	DNA repair protein RadA	2	5	4
Trial 2				
P0ABD8	**Biotin carboxyl carrier protein of acetyl-CoA carboxylase**	**3**	**24**	**33**
C4ZUG5	**50S ribosomal protein L14**	**2**	**21**	**11**
P32132	GTP-binding protein TypA/BipA	3	6	8
P0AE06	**Acriflavine resistance protein A**	**2**	**8**	**6**
P0ABJ9	Cytochrome d ubiquinol oxidase subunit 1	2	5	5
C4ZSQ8	**Ribosome-binding factor A**	**2**	**14**	**4**

Note: Bolded texts represented common proteins in both immunoprecipitation trials

are potential candidates of biotinylated proteins *in vivo*. The four proteins are: the BCCP subunit of acetyl-CoA carboxylase, 50S ribosomal protein L14, acriflavine resistance protein A, and ribosome binding factor A. The identification of BCCP confirms the immunoprecipitation procedures.

Despite the poor reproducibility, we have further characterized the eluted proteins from trial 1 via DAVID and STRING analyses to identify pathways and interacting proteins. **Figure A2-10** shows the protein-protein interaction network (STRING) and **Table A2-2** lists the pathway analysis (DAVID).

A2.3.4. The Confirmation of the ⊿aat Deletion Strain

The use of ∆*aat* cell lysates in the identification of potential *in vivo* L/F transferase substrates ensures that the potential substrates have not been modified by endogenous L/F transferase. To confirm that the ∆*aat* deletion strain indeed has *aat* gene interrupted, I performed whole-cell PCR using primers that flank ~80 bp outside of the 5' and 3' end of *aat* gene on wild-type K-12 cells and ∆*aat* cells. **Figure A2-11** shows the agarose gel electrophoresis analysis of the PCR amplification. The large band at 865 bp resulted from the wild-type K-12 cells corresponds to the length of the *aat* gene (705 bp + primers 160 bp = 865 bp). Meanwhile the band at 2465 bp resulted from the ∆*aat* cells is larger than the *aat* gene, this confirms that the *aat* gene has been interrupted by the kanamycin resistance cassette (*aat* 706 bp + KanR 1600bp + primers 160 bp = 2465

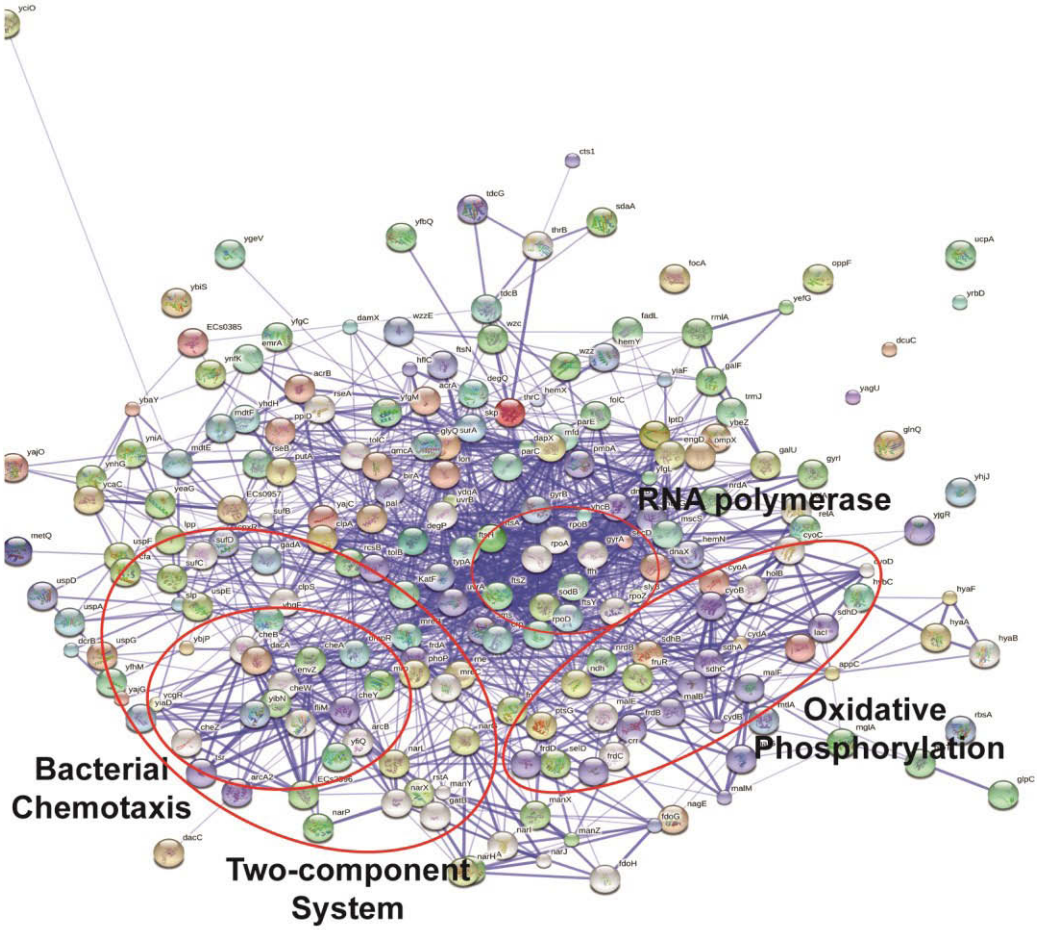

Figure A2-10: A schematic of protein interaction network (STRING analysis) from the elutions of trial 1 show that potential biotinylated proteins are involved in metabolism.

Table A2-2: List of pathways from the elutions by DAVID analysis.

Pathway	Count	%	P-value
Two-component System	17	7.1	7.9e-14
Oxidative Phosphorylation	8	3.4	1.7e-7
Amino Sugar and Nucleotide Sugar Metabolism	6	2.5	4.7e-5
Phosphotransferase System	6	2.5	6.0e-5
Protein Export	4	1.7	5.5e-5
Bacterial Chemotaxis	4	1.7	1.1e-3
Bacterial Secretion System	4	1.7	2.4e-3
Butanoate Metabolism	4	1.7	4.0e-3
Gly, Ser, Thr Metabolism	4	1.7	4.0e-3
Starch and Sugar Metabolism	4	1.7	4.0e-3
Benzoate Degradation via CoA Ligation	3	1.3	9.0e-3
ABC Transporters	7	2.9	9.5e-3
Propanoate Metabolism	3	1.3	2.5e-2
Citric Acid Cycle	3	1.3	3.6e-2
Galactose Metabolism	3	1.3	4.6e-2
Purine Metabolism	3	1.3	4.6e-2
Fructose and Mannose Metabolism	3	1.3	4.6e-2
Biotin Metabolism	2	0.8	6.1e-2
Pyruvate Metabolism	3	1.3	6.5e-2
Trinitrotoluene Degradation	2	0.8	7.1e-2
Nucleotide Excision Repair	2	0.8	7.1e-2

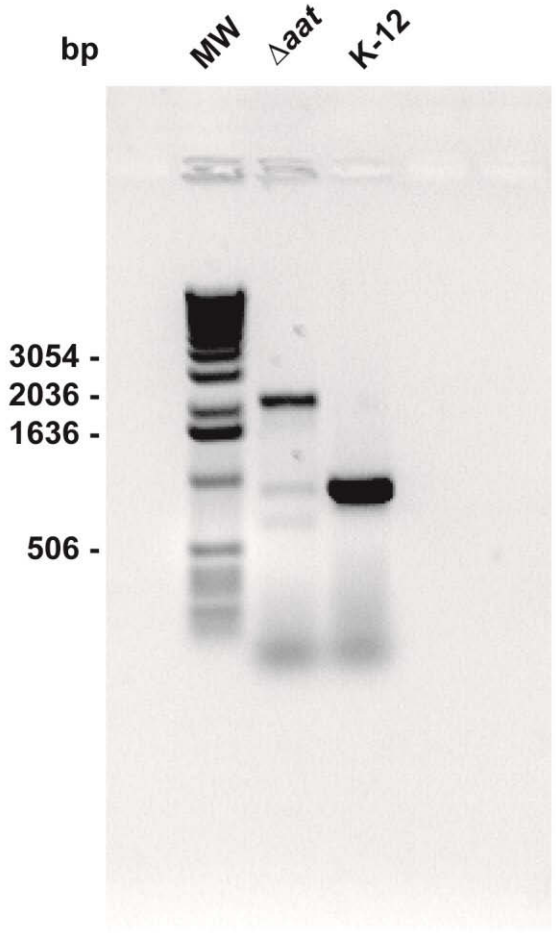

Figure A2-11: 0.8% Agarose gel electrophoresis of wild-type K-12 and Δ*aat* whole cell PCR.

bp). The bands were excise, gel extracted, and sequenced. Purified DNA from the bands confirmed that wild-type K-12 contains the *aat* gene, while the Δ*aat* strain observed an insertion of Tn5 (KanR) gene at position 103 in the positive orientation confirming the disruption of the *aat* gene.

A2.3.5. Δaat Growth Phenotype during Stress

A significant challenge to the identification of *in vivo* substrates of L/F transferase is the lack of knowledge in the cellular events responsible for the generation of L/F transferase substrates. There is a possibility that unnatural amino acid labelling may prove to be not feasible with lysates despite our successful labelling of model polypeptides due to unforeseen challenges. Soffer and Savage suggested that L/F transferase is involved in growth regulatory mechanism (Soffer and Savage 1974). Thus as a preliminary investigation, I have monitored the growth rates of wild-type K-12 and Δ*aat* cells grown under different growth conditions. **Figure A2-12** shows the growth curves ($O.D._{600nm}$ against time) of K-12 and Δ*aat* in rich LB or minimal M9 media grown at normal 37 ºC or heated 42 ºC. There are no significant differences between K-12 and Δ*aat* in rich LB media at either temperature. Although there are differences between the two growth temperatures in minimal M9 media, however there are no significant differences between the growth curves of wild-type K-12 and Δ*aat*. I have additionally tested whether solid media growth alters the growth phenotype. **Figure A2-13** shows the colony growth for wild-type K-12 and Δ*aat* in LB agar or M9 minimal agar after serial dilutions.

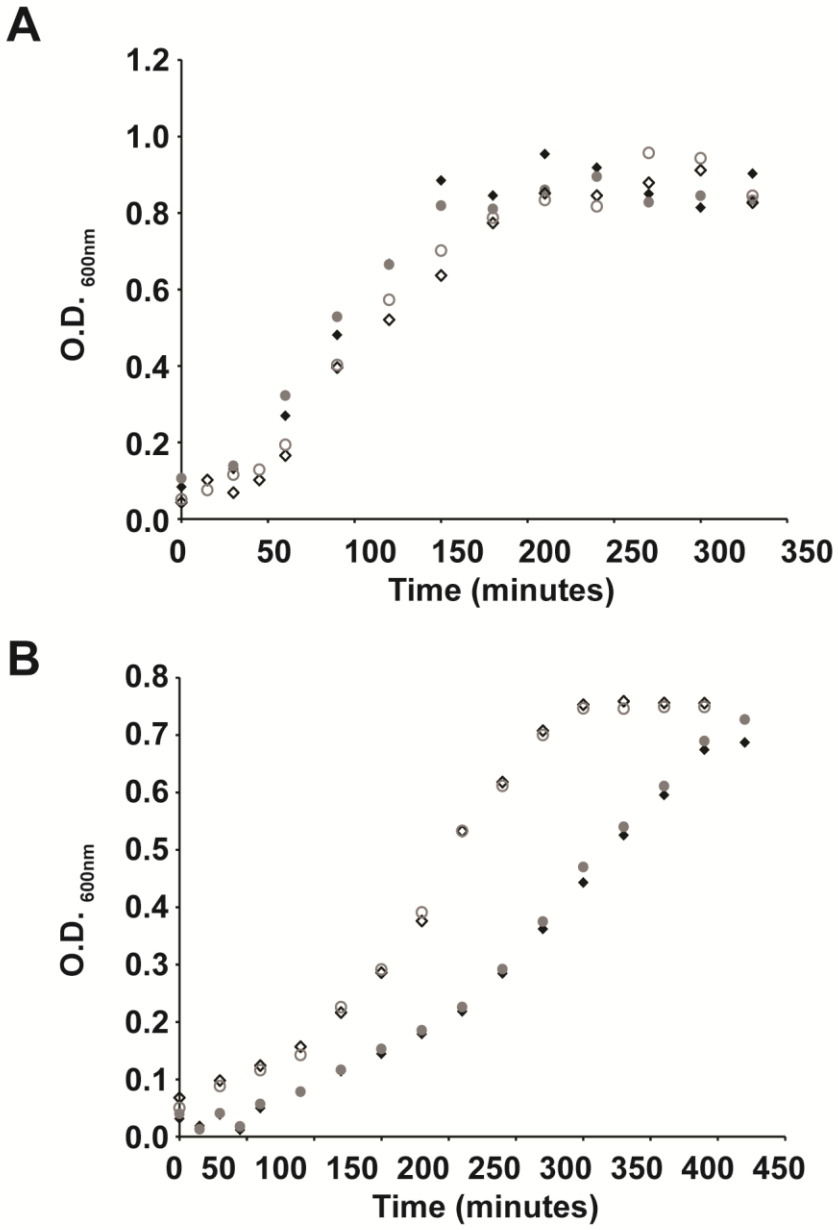

Figure A2-12: Growth curves of wild-type K-12 (black diamonds) and Δ*aat* (grey circles) at 37ºC (filled) and 42ºC (open) in A) rich LB media or B) M9 minimal media.

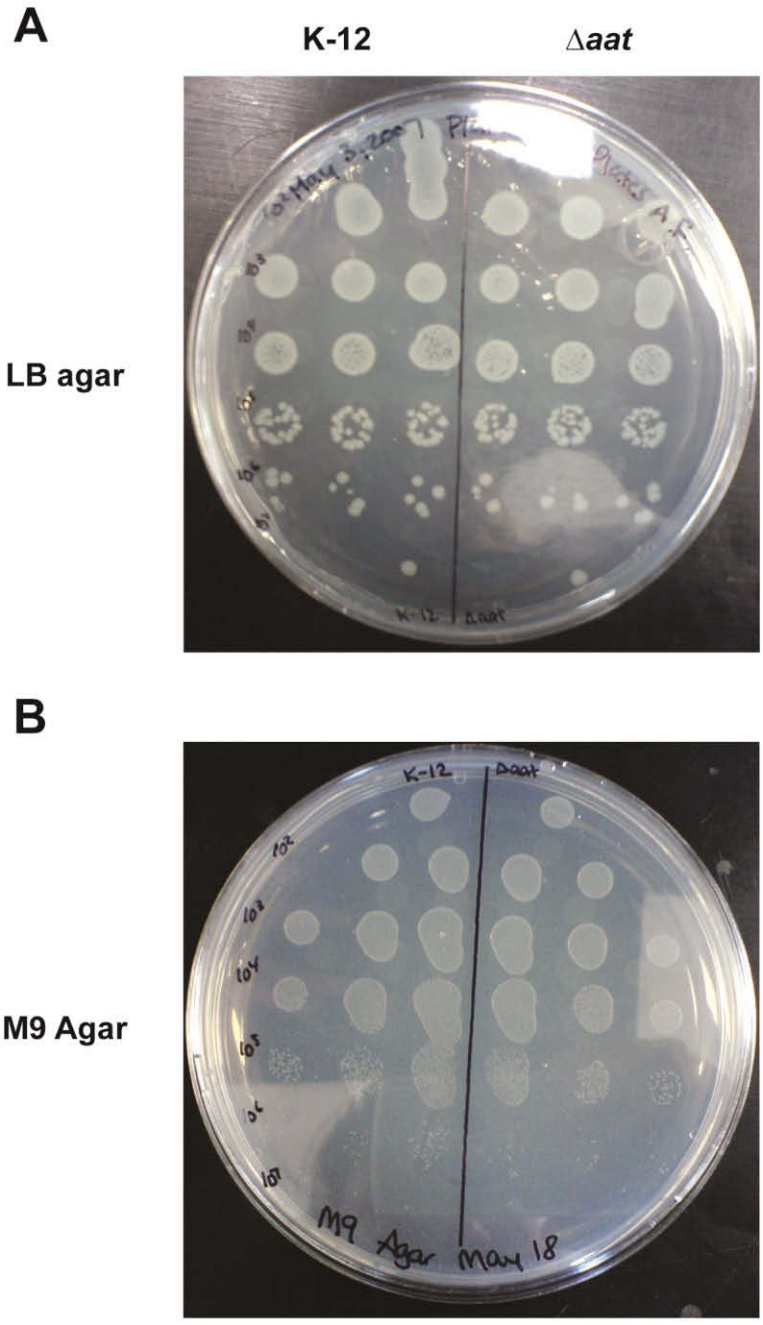

Figure A2-13: Colony morphology of wild-type K-12 (left) and Δ*aat* (right) in (A) rich LB agar or (B) M9 minimal agar in serial dilutions from 10^1 to 10^7.

Despite differences between media conditions, K-12 and Δ*aat* have no significant growth differences in solid phase either. Further experiments with the cells under cold shock (31 ºC) or higher heat shock (45 ºC) for 10 minutes, also does not show differences between wild-type and Δ*aat* (data not shown). Under the conditions tested, I was unable to determine Δ*aat* phenotype.

A2.4. Discussion

A2.4.1. Unnatural Amino Acid Labeling and Click Chemistry Procedures

Although I have showed that the unnatural amino acid labeling on model substrate peptide by L/F transferase and the biotin derivitization via click chemistry were possible, we noticed that a large amount of substrate peptide were not quantitatively reacted. This is possibly due to substrate peptide remains unreacted due to the lack of reactivity of azido-Phe (resonance structures stabilizes the chemical). Additionally based on the observations from **Chapter 5**, we suggest that using a leucine-like unnatural amino acid may significantly enhance the yield of the reactions than phenylalanine-like unnatural amino acids. To overcome the quantitative labeling and click chemistry reactions, we propose to modify the procedure to involve an alkyne-derivative of leucine and an azido-biotin derivative. We have found a commercially available alkyne-derivative of leucine – propargylglycine (Sigma-Aldrich). Tang *et al*. showed that a single mutation of T252Y to LeuRS enables the quantitative

aminoacylation of propargylglycine onto $tRNA^{Leu}$ (Tang and Tirrell 2002). Azido-biotin is also commercially available from several sources.

After optimization with the labeling and click chemistry procedures, we propose to optimize the labeling and click chemistry procedures with casein peptide mixtures. Casein has been previously used as a model substrate in L/F transferase reactions (Leibowitz and Soffer 1971a, Abramochkin and Shrader 1996, Ichetovkin *et al.* 1997, Suto *et al.* 2006, Watanabe *et al.* 2007, Ninnis *et al.* 2009), thus the lysates spiked with a small amount of casein will serve as an internal positive control of the labeling and derivitization procedures. Subsequently the procedures will be optimized for Δ*aat* lysates.

A2.4.2. Streptavidin Immunoprecipitation

The four proteins identified to be potential additional biotinylated proteins are the BCCP subunit of acetyl-CoA carboxylase, 50S ribosomal protein L14, acriflavine resistance protein A, and ribosome binding factor A. The identification of BCCP subunit of acetyl-CoA carboxylase confirms the immunoprecipitation procedures. The remaining three proteins have not been identified to be biotinylated before. Replicates of the streptavidin immunoprecipitation will need to be repeated to ensure the reproducibility of the identification of background, additional biotinylated proteins.

Upon labeling and derivitization of Δ*aat* lysates, potential *in vivo* L/F transferase substrate proteins can be identified through streptavidin

immunoprecipitation for enrichment and LC-MS/MS for identification. The stability of the identified potential protein substrates will be validated *in vivo* in both wild type and *Δaat* cells using Western Blot. L/F transferase substrates are predicted to exhibit significantly shorter half-lives in wild type cells, and their stability will be enhanced in the *Δaat* mutant. Other mutants involved in the N-end rule pathway ()such as *ΔclpS, ΔclpA, ΔclpP* obtained from *E. coli* Keio knockout collection (Baba *et al.* 2006) can be used to confirm the identified substrates as *bona fide* L/F transferase and N-end rule substrates. Additionally, the protein half-life validation method depends heavily on available primary antibodies against bacterial proteins. However, there are limited bacterial primary antibodies available commercially. To overcome this, C-terminal GFP fusion constructs of proteins that have been identified as significant can be obtained from the ASKA (+) clone collection commercially (Kitagawa *et al.* 2005). Anti-GFP antibody is readily available commercially and can be used for Western Blotting.

A2.4.3. Δaat Growth Phenotype

We have also tested whether Δ*aat* have different growth phenotype under nutritional and temperature stress compared to wild-type K-12. We found that under 37 and 42 ºC (or cold shock in 31 ºC and heat shock in 45 ºC) in rich LB media or minimal M9 media, K-12 and Δ*aat* grow with similar profiles. We also observed that there is no significant difference under solid agar growth in neither rich nor minimal media. The negative

identification of a Δ*aat* phenotype may suggest other stress events to trigger eubacterial N-end rule such as oxidation, metal, and more. Alternatively, this may also points to a role for L/F transferase during bacterial infection, since the laboratory strains we worked have deleted infection capabilities.

A2.4.4. Proteome wide identification of ClpS interacting proteins

Recently, over 100 putative *E. coli* N-end rule substrates were identified using ClpS affinity column under non-denaturing conditions during mid-logarithmic (O.D. 0.7) and stationary phase growth (16 hours) at 37 ºC (Humbard *et al.* 2013). Edman degradation confirms a majority of the ClpS-interacting proteins eluted contain N-degrons and approximately 25 % of ClpS-interacting proteins are modified in the cell by L/F transferase, which is consistent with a previous elementary study (Soffer and Savage 1974, Humbard *et al.* 2013).

Despite differences in the immunoprecipitation methods, growth conditions and time of growth between the three ClpS-interacting protein identification studies (**Table 1-2**), a number of proteins including Dps and PATase were identified in more than one study confirming that they are putative *E. coli* N-end rule substrates (Ninnis *et al.* 2009, Schmidt *et al.* 2009, Humbard *et al.* 2013). Many of the identified putative N-end rule substrates belong to large protein complexes (Humbard *et al.* 2013). *E. coli* N-end rule may play a role in the remodelling or quality control of protein complexes. Additionally, certain N-end rule substrates are

observed to be enriched during exponential growth while others are enriched during stationary phase, suggesting that the N-end rule is a widespread and general mechanism of proteolytic processing under different growth conditions in *E. coli*. Therefore, the *E. coli* N-end rule plays a more central role in biological processes than previously recognized including cell division, DNA replication, transcription, translation, metabolism, and protein quality control (Humbard *et al.* 2013).

A2.4.5. Concluding Remarks

Here we have documented the optimization steps to click chemistry coupled LC-MS/MS in the identification of putative L/F transferase *in vivo* substrates. The steps include the unnatural amino acid labeling, click chemistry procedure, enrichment using streptavidin beads, negative control biotinylated protein background identification, and growth conditions. Future experiments will aim to identify putative L/F transferase *in vivo* substrates under different growth conditions and can be compared to the substrates identified through ClpS-interacting proteins. The biological functions of *E. coli* N-end rule may be revealed with the identification of *in vivo* substrates.

A2.5. References

Abramochkin, G., and Shrader, T.E. (1996) Aminoacyl-tRNA recognition by the leucyl/phenylalanyl-tRNA-protein transferase. *J.Biol.Chem.* **271,** 22901-22907

Atlas, R.M. (1993) *Handbook of Microbiological Media ,* Second Ed., CRC Press

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K.A., Tomita, M., Wanner, B.L., and Mori, H. (2006) Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol.Syst.Biol.* **2,** 2006.0008

Barker, D.F., and Campbell, A.M. (1980) Use of bio-lac fusion strains to study regulation of biotin biosynthesis in *Escherichia coli. J.Bacteriol.* **143,** 789-800

Chapman-Smith, A., and Cronan, J.E.,Jr (1999) The enzymatic biotinylation of proteins: a post-translational modification of exceptional specificity. *Trends Biochem.Sci.* **24,** 359-363

Connor, R.E., Piatkov, K., Varshavsky, A., and Tirrell, D.A. (2008) Enzymatic N-terminal addition of noncanonical amino acids to peptides and proteins. *Chembiochem.* **9,** 366-369

Ebhardt, H.A., Xu, Z., Fung, A.W., and Fahlman, R.P. (2009) Quantification of the post-translational addition of amino acids to proteins by MALDI-TOF mass spectrometry. *Anal.Chem.* **81,** 1937-1943

Ebisu, K., Tateno, H., Kuroiwa, H., Kawakami, K., Ikeuchi, M., Hirabayashi, J., Sisido, M., and Taki, M. (2009) N-terminal specific point-immobilization of active proteins by the one-pot NEXT-A method. *Chembiochem.* **10,** 2460-2464

Eisenberg, M.A. (1975) Mode of action of alpha-dehydrobiotin, a biotin analogue. *J.Bacteriol.* **123,** 248-254

Erbse, A., Schmidt, R., Bornemann, T., Schneider-Mergener, J., Mogk, A., Zahn, R., Dougan, D.A., and Bukau, B. (2006) ClpS is an essential component of the N-end rule pathway in *Escherichia coli. Nature.* **439,** 753-756

Green, N.M. (1975) Avidin. *Adv.Protein Chem.* **29,** 85-133

Humbard, M.A., Surkov, S., De Donatis, G.M., Jenkins, L.M., and Maurizi, M.R. (2013) The N-degradome of *Escherichia coli*: limited proteolysis *in vivo* generates a large pool of proteins bearing N-degrons. *J.Biol.Chem.* **288,** 28913-28924

Ichetovkin, I.E., Abramochkin, G., and Shrader, T.E. (1997) Substrate recognition by the leucyl/phenylalanyl-tRNA-protein transferase. Conservation within the enzyme family and localization to the trypsin-resistant domain. *J.Biol.Chem.* **272,** 33009-33014

Kitagawa, M., Ara, T., Arifuzzaman, M., Ioka-Nakamichi, T., Inamoto, E., Toyonaga, H., and Mori, H. (2005) Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research. *DNA Res.* **12,** 291-299

Kuno, A., Taki, M., Kaneko, S., Taira, K., and Hasegawa, T. (2003) Leucyl/phenylalanyl (L/F)-tRNA-protein transferase-mediated N-terminal specific labelling of a protein in vitro. *Nucleic Acids Res.Suppl.* **(3),** 259-260

Leibowitz, M.J., and Soffer, R.L. (1969) A soluble enzyme from *Escherichia coli* which catalyzes the transfer of leucine and phenylalanine from tRNA to acceptor proteins. *Biochem.Biophys.Res.Commun.* **36,** 47-53

Leibowitz, M.J., and Soffer, R.L. (1971a) Enzymatic modification of proteins. VII. Substrate specificity of leucyl,phenylalanyl-transfer ribonucleic acid-protein transferase. *J.Biol.Chem.* **246,** 5207-5212

Leibowitz, M.J., and Soffer, R.L. (1971b) Modification of a specific ribosomal protein catalyzed by leucyl, phenylalanyl-tRNA: protein transferase. *Proc.Natl.Acad.Sci.U.S.A.* **68,** 1866-1869

Ninnis, R.L., Spall, S.K., Talbo, G.H., Truscott, K.N., and Dougan, D.A. (2009) Modification of PATase by L/F-transferase generates a ClpS-dependent N-end rule substrate in *Escherichia coli. EMBO J.* **28,** 1732-1744

Rao, P.M., and Kaji, H. (1974) Utilization of isoaccepting leucyl-tRNA in the soluble incorporation system and protein synthesizing systems from *E.coli. FEBS Lett.* **43,** 199-202

Rostovtsev, V.V., Green, L.G., Fokin, V.V., and Sharpless, K.B. (2002) A stepwise huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew.Chem.Int.Ed Engl.* **41,** 2596-2599

Schmidt, R., Zahn, R., Bukau, B., and Mogk, A. (2009) ClpS is the recognition component for *Escherichia coli* substrates of the N-end rule degradation pathway. *Mol.Microbiol.* **72,** 506-517

Soffer, R.L., and Savage, M. (1974) A mutant of *Escherichia coli* defective in leucyl, phenylalanyl-tRNA-protein transferase. *Proc. Natl. Acad. Sci. U.S.A.* **71,** 1004-1007

Speers, A.E., and Cravatt, B.F. (2004) Profiling enzyme activities *in vivo* using click chemistry methods*. Chem.Biol.* **11,** 535-546

Suto, K., Shimizu, Y., Watanabe, K., Ueda, T., Fukai, S., Nureki, O., and Tomita, K. (2006) Crystal structures of leucyl/phenylalanyl-tRNA-protein transferase and its complex with an aminoacyl-tRNA analog*. EMBO J.* **25,** 5942-5950

Taki, M., Kuroiwa, H., and Sisido, M. (2008) Chemoenzymatic transfer of fluorescent non-natural amino acids to the N terminus of a protein/peptide*. Chembiochem.* **9,** 719-722

Taki, M., Kuroiwa, H., and Sisido, M. (2009) The NEXT-A (N-terminal EXtension with Transferase and ARS) reaction*. Nucleic Acids Symp.Ser.(Oxf).* **(53),** 37-38

Taki, M., and Sisido, M. (2007) Leucyl/phenylalanyl(L/F)-tRNA-protein transferase-mediated aminoacyl transfer of a nonnatural amino acid to the N-terminus of peptides and proteins and subsequent functionalization by bioorthogonal reactions*. Biopolymers.* **88,** 263-271

Tang, Y., and Tirrell, D.A. (2002) Attenuation of the editing activity of the *Escherichia coli* leucyl-tRNA synthetase allows incorporation of novel amino acids into proteins in vivo*. Biochemistry.* **41,** 10635-10645

Tornoe, C.W., Christensen, C., and Meldal, M. (2002) Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides*. J.Org.Chem.* **67,** 3057-3064

Wagner, A.M., Fegley, M.W., Warner, J.B., Grindley, C.L., Marotta, N.P., and Petersson, E.J. (2011) N-terminal protein modification using simple aminoacyl transferase substrates*. J.Am.Chem.Soc.* **133,** 15139-15147

Wang, Q., Chan, T.R., Hilgraf, R., Fokin, V.V., Sharpless, K.B., and Finn, M.G. (2003) Bioconjugation by copper(I)-catalyzed azide-alkyne [3 + 2] cycloaddition*. J.Am.Chem.Soc.* **125,** 3192-3193

Watanabe, K., Toh, Y., Suto, K., Shimizu, Y., Oka, N., Wada, T., and Tomita, K. (2007) Protein-based peptide-bond formation by aminoacyl-tRNA protein transferase*. Nature.* **449,** 867-871